DOCUMENT RESUME

ED 386 494 TM 024 063

AUTHOR	Lawrence, Ida M.
TITLE	The Effect of Test Speededness on Subgroup Performance.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-93-49
PUB DATE	Oct 93
NOTE	71p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Asian Americans; Black Students; *College Entrance Examinations; *Difficulty Level; Females; Higher Education; High Schools; Hispanic Americans; *Mathematics Tests; *Racial Differences; *Test Items
IDENTIFIERS	Graduate Management Admission Test; *Scholastic Aptitude Test; *Speededness (Tests)

ABSTRACT

The purpose of this study was to gauge the disadvantage, if any, of reaching difficult items that appear earlier in a test, at the expense of not reaching easier items that appear later in a test. The study focused on Scholastic Aptitude Test (SAT) and Graduate Management Admissions Test (GMAT) data and examined the effect for females, Asian Americans, Blacks, and Hispanics. The impact of differential test speededness on subgroup differences in test scores was found to be minimal for SAT Mathematical and GMAT Quantitative tests. However, there appears to be a relationship between the presence of differential speededness on the SAT Verbal test and subgroup differences in test performance; after matching on total score, Black and Hispanic examinees receive credit for more difficult sets of items in the SAT Verbal test, relative to White examinees. Due to smaller sample sizes, findings for the GMAT Verbal tests were not as clear-cut as the findings for the SAT Verbal tests. This research attempted to provide a technique for assessing the circumstances under which the effect of test speededness has a differential effect on the performance of subgroups, when differences in ability are taken into consideration. Results of this study cannot predict how examinees would have performed with a shorter test or more liberal time limits, because test strategies might have been different under those conditions. Eighteen tables and 24 figures present analysis data. (Contains 8 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

RR-93-49

ED 386 494

RESEARCH REPORT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. I. BRAUN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

THE EFFECT OF TEST SPEEDEDNESS ON SUBGROUP PERFORMANCE

Ida M. Lawrence

Educational Testing Service
Princeton, New Jersey
October 1993

TM024063

BEST COPY AVAILABLE

THE EFFECT OF TEST SPEEDEDNESS ON SUBGROUP PERFORMANCE

Ida M. Lawrence[1]

[1]I thank Karen Carroll for her programming efforts, Bernadette Stephens for producing the graphics, and Wendi Wright and Doris Larsen for their help in obtaining and preparing the data. I thank Gerald Melican, Skip Livingston, Alicia Schmitt, Larry Wightman, and Nancy Wright for helpful comments on an earlier draft of this paper. I especially thank Neil Dorans for his help in planning the study and reviewing the report. Support for this research was provided by Program Research Planning Council.

Copyright © 1993. Educational Testing Service. All rights reserved.

Abstract

The purpose of this study was to gauge the disadvantage, if any, of reaching difficult items that appear earlier in a test, at the expense of not reaching easier items that appear later in a test. The study focused on SAT and GMAT data and examined the effect for the following focal groups: females, Asian American, Blacks, and Hispanics. The impact of differential test speededness on subgroup differences in test scores was found to be minimal for SAT Mathematical and GMAT Quantitative tests. However, there appears to be a relationship between the presence of differential speededness on the SAT Verbal test and subgroup differences in test performance; after matching on total score, Black and Hispanic examinees receive credit for more difficult sets of items in the SAT Verbal test, relative to White examinees. Due to smaller sample sizes, findings for the GMAT Verbal tests were not as clear-cut as the findings for the SAT Verbal tests. This research attempted to provide a technique for assessing the circumstances under which the effect of test speededness has a differential effect on the performance of subgroups, when differences in ability are taken into consideration. As is the case with all non-experimental research on test speededness, the results of this study cannot predict how examinees would have performed with a shorter test, or more liberal time limits, because test strategies might have been different under those conditions.

THE EFFECT OF TEST SPEEDEDNESS ON SUBGROUP PERFORMANCE

As part of an effort to evaluate test speededness, one area of concern is the degree to which subgroup performance is differentially affected by not completing a test section. A difference in scores between subgroups may occur when candidates in particular subgroups tend not to reach the relatively easy items appearing at the end of a test section. The purpose of this study is to gauge the disadvantage, if any, of reaching the difficult items that appear earlier in the test, at the expense of not reaching the easier items that appear later in the test.

When a test is speeded, it is desirable to order items within a test according to their difficulty, with the easier items appearing at the beginning of the test and the harder items appearing at the end of the test. Under this kind of item ordering, students who are unable to complete the test will at least have an opportunity to attempt those items for which the probability of a correct response is highest (assuming that students pace through the items in the order presented). For this reason, a typical recommendation for test construction is that test items in sections be arranged in ascending order of difficulty (e.g., Cronbach, 1951).

Sometimes it is not feasible to order items from easy to hard. An example of a test in which the items cannot all be ordered from easy to hard is a test that contains passage-based reading comprehension items. In general, reading questions do not progress from easy to difficult. Instead, the questions follow the logic and organization of the passages. When tests are speeded, the effect on test performance of this kind of item ordering may differ for subgroups, and this difference raises an equity issue. This situation would occur when candidates from a particular subgroup systematically tend not to have sufficient time to attempt the relatively easy items at the end of the test section. This phenomenon

was shown at the item level by Schmitt and Bleistein (1987) and Dorans, Schmitt and Curley (1988). In both of these studies, Black students were found to reach fewer items at the end of a SAT verbal section when compared to a matched group of White students.

Data from one of the forms analyzed in the present study (Form F) provide an example of this situation. The difference between matched groups of Blacks and Whites in the percentage of examinees not reaching the last five items ranged between 2% and 7% for one of the verbal sections, and ranged between 8% and 12% for the other verbal section. Some of the items not reached by Black examinees at the end of the test section were fairly easy (as indicated in Table 2a and 2b of this paper, which shows the percent correct for each item in the test). This set of results is typical for SAT verbal sections.

The purpose of this study was to see if it can be demonstrated empirically that speeded tests where items are not ordered with respect to difficulty have a differential effect on subgroup performance. Examples of these kinds of tests are found in the Scholastic Aptitude Test (SAT) and Graduate Management Aptitude Test (GMAT) programs. Data from national administrations indicate that both of these formula-scored tests are speeded, and that the degree of speededness varies by subgroup.

The GMAT Reading Comprehension section is composed of 23 or 25 items based on three reading passages; relatively easy items are associated with all three passages, thus, easy items appear at the beginning, middle, and end of the section. The verbal portion of the Scholastic Aptitude Test (SAT) is another example of a test where some items are not ordered from easy to hard. The SAT verbal test is composed of three discrete item types, as well as reading comprehension items, which are passage-based. Although the discrete items are generally ordered from easy to hard within item type, the reading

comprehension items associated with separate passages cannot ordered from easy to hard. Consequently, some of the items positioned toward the end of the SAT verbal sections are relatively easy items. For both of these testing programs, items in the Mathematical (SAT) or Quantitative (GMAT) tests are generally ordered from easy to hard.

At the end of their literature review of the effects of item arrangement, Leary and Dorans (1985) ask the question, "Are there identifiable subgroups of examinees who are more affected by variation in item position?" (p. 411). This research was an attempt to answer this question under speeded conditions. The present research is also related to the role of context effects, as described by Brennan (1992). If the placement of easy items at the end of a speeded test section has a systematic effect on test performance for certain subgroups relative to others, one could argue that an item position context effect is at play.

In sum, the objective of this study was to provide information concerning the effect of differential test speededness on subgroup performance on the SAT and GMAT. The information is provided through the use a scoring rule that takes item difficulty into account. The assumption in this research is that slower examinees, who tend not to reach the easier items at the end of the test, would receive higher test scores under a scoring rule that weights item scores by item difficulty, relative to faster examinees with the same formula score. While a scoring rule that weights item scores by difficulty level is <u>not</u> recommended for operational use, it provides an opportunity to analyze performance data in an informative manner, and may indicate that the configuration of items in the SAT and GMAT underestimates the ability of slower examinees.

Data

The data source for this study was test data from national administrations of the SAT and GMAT. Analyses were carried out on seven forms of the SAT

Verbal test, six forms of the SAT Mathematical test, three forms of the GMAT Verbal test and three forms of the GMAT Quantitative test. All of these tests are formula scored to correct for guessing (rights minus one-quarter wrongs for five-choice items and rights minus one-third wrongs for four-choice items).

Table 1 shows the test forms analyzed and the associated sample sizes. Note that the samples represent a variety of administration times for the SAT and GMAT. The samples for the SAT analyses were restricted to examinees who indicated on their registration form that they were high school juniors or seniors and that English was at least one of their first spoken languages. The samples for the GMAT analyses were restricted to examinees who indicated on the answer sheet that English was their best language and indicated on the registration form that they were citizens of the United States. In addition, only those examinees who classified themselves with respect to race or ethnicity were included in the samples for the GMAT analyses.

Methodology

For each test form, within each subgroup, samples were stratified with respect to total formula score.[1] The following subgroups were identified: Female, Male, Asian American, Black, Hispanic, and White students. For comparisons based on ethnicity/race, White students served as the reference group; for comparisons based on gender, male students served as the reference group. Within each total score stratum, a determination was made of the percentage of examinees within each subgroup who:

(a) answered the item correctly	R
(b) answered the item incorrectly	W
(c) omitted the item	O

[1]Reliability estimates (internal consistency) for the matching criterion range between .90 and .93 for SAT Verbal, .90 and .92 for SAT Mathematical, .86 and .88 for GMAT Verbal and .91 and .92 for GMAT Quantitative.

(d) did not reach the item **NR**

When a test taker does not respond to an item, but responds to subsequent items, the non response to that item is referred to as an omit (O). If the test taker does not respond to any of the subsequent items in a separately timed section of the test, then the first non response and the subsequent non responses are referred to as "not reached" (NR).

Comparative Data on Test Speededness

Because the focus of this research was on the effect of test speededness, several approaches to comparing speededness across subgroups were used. First the average number of items not reached on the test was compared. Second, plots of the average number of items not reached on the test, conditioned on formula score, were compared. Third, indices of differential speededness for focal and reference groups, matched on formula scores, were computed. The latter indices were based on formulas 1a and 1b, below.

Formula 1.a -- Conditional Differences in Not Reaching

$$CDNR_m = \Sigma_i \ (PNR_{fim} - PNR_{rim}) \times NI$$

Formula 1.b -- Average Conditional Differences in Not Reaching

$$ACDNR = \Sigma_m \ [F_{fm} \ / \ \Sigma F_{fm}] * CDNR_m$$

Where, PNR_{fim} and PNR_{rim} are the proportions **not reaching** item **i** in the **focal** and **reference** groups at score level **m**,

NI is the number of items in the test

F_{fm} is the number of **focal** group examinees at score level **m**.

Formula 1a provides an index of focal versus reference group average number of items not reached at the end of the test section, conditioned on total score. Formula 1b provides an index that summarizes the differentials in average number of items not reached across score levels, for an overall index of focal versus reference groups differential not reaching. This index is in the metric of the number of items in the test, thus the index indicates the average difference in the number of items not reached by the focal group relative to the reference group, after matching the groups on total score.

Differences in Difficulty Level of Items Contributing to Formula Score

The following indices were used to assess whether the focal group and reference group differed, on average, in terms of the difficulty level of items that contributed to their formula scores.

Formula 2.a -- Weighted Conditional Difference Score ($WCDS_m$)

$$WCDS_m = \frac{\sum_i \left[\left(PR_{fim} - PR_{rim} \right) - \left(\left(PW_{fim} - PW_{rim} \right) \left(\frac{1}{k_i - 1} \right) \right) \right] * W_i}{\sum W_i} * NI$$

Formula 2.b – Average Weighted Conditional Difference Score (AWCDS)

$$AWCDS = \sum_m \left[F_{fm} / \sum F_{fm} \right] * WCD_m$$

where, PR_{fim} and PR_{rim} are the proportions **Right** in the **focal** and **reference** groups on item **i** at score level **m**,

PW_{fim} and PW_{rim} are the proportions **Wrong** in the **focal** and **reference** groups on item **i** at score level **m**,

k_i is the number of options in item **i**, and

$W_i = (1-p_i)$, where p is the proportion correct on item **i** for the total group (based on item analysis sample)[2]

NI is the number of items in the test

F_{fm} is the number of **focal** group examinees at score level **m**.

Formula 2a makes use of weights (W_i) to express item difficulty. The weights chosen for this study were 1 minus the percent correct based on the item analysis sample for each item in the test. Several other weights could have been chosen. For example, a variety of strategies for weighting item scores and item responses are reviewed by Wang and Stanley (1970). However, the purpose of item weighting in the studies reviewed by Wang and Stanley was to achieve gains in reliability or validity; the purpose of item weighting in the present study was to characterize the effect of test speededness on test performance. The choice of $W_i = 1-p_i$ provided a convenient way to weight differences in item scores by item difficulty.

Formulas 2.a and 2.b indicate differences in focal group and reference group formula scores for examinees who are matched with respect to formula score. Because the groups are matched on formula score, the conditional difference between the groups reduces to zero[3] when w_i equals 1 for all items,

---

[2]Students who did not reach the item were not included in the computation of item difficulty.

[3]Actually, the difference is only approximately zero because the groups are matched on integer formula scores and the computed differences are expressed as unrounded formula scores.

i.e., all differences in item responses are given equal weight, which is the weight used when SAT and GMAT scores are reported.

When W_i are estimates of item difficulty, $WCDS_m$ and AWCDS will be non-zero. A non-zero value indicates that the focal and reference groups received equivalent average formula scores on the test but that the items contributing to the formula scores differ in terms of difficulty. The sign of $WCDS_m$ will be positive if the average formula score for the focal group is based on more difficult items than those taken by the the reference group, and will be negative if the average formula score for the focal group is based on easier items than those taken by the reference group. AWCDS is an index of differences averaged over score levels; it is reported in the metric of formula score units on the total test.

Note that a non-zero value of AWCDS can occur under two conditions:
Condition A -- The reference group and the focal groups are attempting items that differ in difficulty;
Condition B -- The reference group and the focal group respond to the same set of items but one group tends to answer some difficult items correctly while missing some easy ones, and the other group tends to answer some easy items correctly while missing some difficult ones.

In general, the items in the tests in this study do not exhibit differential item functioning (DIF). This means that the item characteristic curves (ICCs) for the focal group and reference group are essentially equivalent. Therefore, at a given formula score, the probability of a correct response is similar in the focal group and the reference group. Furthermore, the ICCs for these items tend to be monotonically increasing functions. Given minimal DIF on these items, and monotonically increasing ICCs, it is unlikely that non-zero values of AWCDS would occur because examinees from the focal group tend to answer difficult

items correctly and miss easy ones. It is more likely that examinees from the focal group and the reference group tend to respond to different items in the test. The rationale behind the use of formulas 2a and 2b is that, under these circumstances, Condition A is more likely than Condition B.

Differences in Omitting Patterns for Subgroups

After looking at the findings on subgroup differences in the difficulty level of attempted items, it became apparent that for some focal group/reference group comparisons, differential rates of omitting an item, in addition to differential rates of not reaching an item, needed to be examined as an aid to interpreting the results. Therefore, the following indices were used to assess differential rates of omitting for focal and reference groups, matched on formula scores.

Formula 3.a -- Conditional Difference in Omitting

$$CDO_m = \Sigma_i (PO_{fim} - PO_{rim}) \times NI$$

Formula 3.b -- Average Conditional Difference in Omitting

$$ACDO = \Sigma_m [F_{fm} / \Sigma F_{fm}] * CDO_m$$

Where, PO_{fim} and PO_{rim} are the proportions **omitting** item i in the **focal** and **reference** groups at score level m,

NI is the number of items in the test

F_{fm} is the number of **focal** group examinees at score level m.

Formula 3a provides an index of focal versus reference group rates of omitting, conditioned on total score. Formula 3b provides an index that summarizes the omitting rate differentials across score levels, for an overall index of focal versus reference groups omitting. By multiplying by NI, CDO_m is in the metric of the number of items in the test. ACDO indicates the average difference in the number of items omitted by the focal group relative to the reference group, after matching the groups on total score.

Results

A description of the results of this study is divided into three sections. The first section provides information about the ordering of items in SAT and GMAT separately timed sections. The second section provides information about test speededness for the subgroups. The third section describes subgroup differences in the difficulty level of items contributing to formula scores.

1. Ordering of Items on the SAT and GMAT

Table 2a and Table 2b show the relationship between item difficulty (expressed as percent correct) and item position within the separately timed Verbal 1 and Verbal 2 sections of seven SAT forms. Both sections of the SAT verbal test contain four different item types, configured as follows:

Verbal 1	Items 1 - 15	Antonyms
	Items 16 - 25	Sentence Completion
	Items 26 - 35	Reading Comprehension
	Items 36 - 45	Analogies
Verbal 2	Items 1 - 10	Antonyms
	Items 11 - 15	Sentence Completion

Items 16 - 25 Analogies

Items 26 - 40 Reading Comprehension

The items are ordered from easy to hard within item type (shown within blocks of the table). This relationship between item difficulty and item position for the SAT Verbal sections is depicted graphically in Figure 1 for one of the SAT forms (Form A).

Table 3a and Table 3b show the relationship between item difficulty (expressed as percent correct) and item position within the separately timed Math 1 and Math 2 sections of six SAT forms. The Math 1 section contains only one item type (5-choice multiple choice items). For this section, the items are generally arranged from easiest to hardest. The Math 2 section contains two item types (shown within blocks of the table), with regular math items appearing at the beginning and the end of the section and quantitative comparison items appearing in the middle of the section. The SAT Mathematical sections are configured as follows:

Math 1	Items 1 - 25	5-choice multiple-choice
Math 2	Items 1 - 7	5-choice multiple-choice
	Items 8 - 27	4-choice quantitative comparison
	Items 28 - 35	5-choice multiple-choice

The shift in item types accounts for the fact that some easy items appear toward the end of the Math 2 section. The relationship between item difficulty and item position for the SAT Mathematical sections is depicted graphically in Figure 2 for one of the SAT forms (Form A).

Table 4 shows the relationship between item difficulty and item position for the three separately timed sections contributing to the Verbal score on the GMAT. Items in the Critical Reasoning section are generally arranged from easiest to hardest. For the other two sections (Reading Comprehension and Sentence Correction), several items at the end of the sections are relatively easy items. The relationship between item difficulty and item position for the GMAT Verbal sections is depicted graphically in Figure 3 for one of the GMAT forms (Form X)

Table 5 shows the relationship between item difficulty and item position for the three separately timed sections contributing to the Quantitative score on the GMAT. Similar to the mathematical portion of the SAT, the items in these sections tend to be ordered from easiest to hardest. The relationship between item difficulty and item position for the GMAT Quantitative sections is depicted graphically in Figure 4 for one of the GMAT forms (Form X) .

2. Subgroup Speededness Data on the SAT and GMAT

Table 6 shows the average number of items on SAT Verbal forms not reached by focal and reference groups. In general, the tests are similarly speeded for males and females. The tests are least speeded for White and Asian American examinees, and most speeded for Black and Hispanic examinees. A similar pattern exists for the SAT Mathematical forms (Table 7).

Table 8 shows the average number of items on GMAT Verbal forms not reached by focal and reference groups. The tests are slightly more speeded for females than for males. The tests are most speeded for Black examinees, followed by Hispanic examinees, followed by Asian American examinees. The tests are least speeded for White examinees. The Quantitative tests (Table 9) are

more speeded for females than for males. The tests are most speeded for Black examinees.

The comparative speededness data presented in Tables 6-9 are affected by group ability, where the tests are more speeded for the lower scoring groups. The data shown in Figures 5-23 indicate that the tests are more speeded for certain focal groups relative to the reference group even after conditioning on total test score. The plots display the average number of items not reached for subgroups as a function of formula score level. Each data point is based on a minimum of ten examinees, which is why the range of formula scores on the x-axis varies for each test form. Formula 1b was computed to summarize differences in average conditional not reaching patterns for focal and reference groups. These data are presented in Table 10 for SAT Verbal tests, Table 11 for SAT Mathematical tests, Table 12 for GMAT Verbal tests, and Table 13 for GMAT Quantitative tests.

SAT Verbal Tests. Figures 5 - 11 indicate that there is a monotonically decreasing relationship, as expected, between the number of items not reached in the tests and formula score. The average number of items not reached is generally similar for males and females. The tests appear to be more speeded (as indicated by a higher average number of items not reached) for Black and Hispanic examinees, relative to similarly performing White examinees, primarily at the lower score levels.[4] The tests are similarly speeded for White and Asian American examinees. Table 10 shows that the SAT Verbal tests are more speeded (higher rates of not reaching) for Black and Hispanic examinees, relative to White

[4]This research assumes that examinees pace through the items in the test and omit when necessary. However, there may be other explanations for why Black examinees leave more items unanswered at the end of the test section than White examinees with the same formula score. An alternative explanation is that While examinees are more inclined to guess at the difficult items at the ends of the sections. Another possibility is that Black examinees are less test-wise with regard to answering easy items first.

examinees, after matching on total score. These effects are consistent across test forms.

SAT Mathematical Tests. Figures 12 - 17 indicate that the relationship between the number of items not reached in the tests and formula score is curvilinear. Contrary to the pattern seen for the SAT Verbal forms, it appears that very low scoring examinees tend to attempt the difficult items at the end of the SAT Mathematical sections rather than decide to leave them blank. The tests are similarly speeded for males and females. The tests are slightly more speeded for Black and Hispanic examinees, relative to similarly performing White and Asian American examinees, and this effect spans most of the score range. Table 11 shows that the SAT mathematical tests are more speeded (higher rates of not reaching) for Black and Hispanic examinees, relative to White examinees, after matching on total score. These effects are consistent across test forms.

GMAT Verbal Tests. Figures 18 - 20 indicate that there is a monotonically decreasing relationship between the number of items not reached in the tests and formula score. The average number of items not reached is generally similar for males and females. It is difficult to evaluate the comparisons based on ethnic/race subgroups due to scarceness of data, although the tests appear to be slightly more speeded for Black examinees relative to White examinees. Table 12 shows that the tests are consistently more speeded for Black examinees relative to White examinees.

GMAT Quantitative Tests. Figures 21 - 23 indicate that the relationship between the number of items not reached in the tests and formula score is linear from formula score 20 and up. Similar to the phenomenon observed for SAT Mathematical, very low scoring examinees tend to attempt the difficult items at the end of the test sections rather than decide to leave them blank. The average number of items not reached is generally similar for males and females. It is

difficult to evaluate the comparisons based on ethnic/race subgroups due to scarceness of data. However, the tests appear to be more speeded for Black examinees relative to White examinees. Table 13 shows that the tests are consistently more speeded for Black examinees relative to White examinees. The data for Hispanic examinees and Asian American examinees are not as consistent as for the Black examinees, but there is a tendency for the GMAT Quantitative tests to also be more speeded for Hispanic examinees, relative to matched groups of White examinees. The inconsistency of findings across test forms may be a consequence of relatively smaller sample sizes.

3. Effects of Weighting Item Responses by Difficulty

Tables 14 - 17 show the effects on differences in subgroup performance of weighting item responses by item difficulty. A positive value of AWCDS indicates that the focal group received formula score credit for a more difficult set of items than the reference group, given the same formula score on the test, and a negative value indicates an easier set of items. The values in these tables indicate the degree (in the metric of raw scores on the test) to which the focal group would be advantaged or disadvantaged, relative to the reference group, if item responses were weighted by item difficulty, giving more weight to the difficult items and less weight to the easier items in the test.

SAT Verbal Tests. The data shown in Table 14 indicate that weighting of item responses by difficulty has little effect on differences in SAT Verbal scores for males and females; the average difference over seven SAT forms is -.02 raw scores. This suggests that males and females receiving the same formula score on the SAT verbal tests do so by receiving credit for items of similar difficulty level. In contrast, Black, Hispanic, and Asian American examinees appear to be earning formula score credit on more difficult items in the test, relative to White

examinees. Weighting item responses by item difficulty would advantage Black and Hispanic examinees by approximately 4-tenths of a raw score point on the average and would advantage Asian American examinees by approximately seven-tenths of a raw score. A difference larger than .5 raw scores is notable because a raw score difference of .5 or more would translate into at least a 10 point difference in reported scores on the 200-to-800 scale, which is reported in increments of 10. Thus, the observed differences would have an effect on examinees reported scores.

The findings for Blacks and Hispanics are not surprising, given the fact that the SAT Verbal test is more speeded for these subgroups, relative to White examinees of similar ability. However, the results for Asian American examinees, where the largest effect was found, was not anticipated because the SAT verbal test is generally not speeded for this subgroup, relative to White examinees (see Table 6). A follow-up analysis, looking at omitting patterns, was carried out, and the data are shown in Table 18. Formula 3b was used to compute the indices displayed in this table. Note that while Asian American examinees and White examinees do not differ with respect to test speededness, Asian American examinees do tend to omit more items than White examinees of similar ability. In fact, with the exception of Black examinees, Asian American examinees have the highest rate of omitting. Coupled with the finding that Asian American examinees tend to receive formula score credit for a more difficult set of items than White examinees, this finding implies that Asian American examinees are tending to omit more easy items than White examinees of similar ability.

This possibility was investigated with data from one of the forms in this study. Form SAT E was selected because the value of AWCDS was the highest and the rate of differential omitting was also the highest. Figure 24 shows the

relationship between differential omitting for Asian Americans on each item in the verbal test of Form SAT E and item difficulty. In general, Asian American examinees omit items at a higher rate than White examinees of comparable ability; some of these items are fairly easy (percent correct above 75%).

SAT Mathematical Tests. The data shown in Table 15 indicate that weighting of item responses by difficulty has little effect on differences in SAT Mathematical scores for males and females; the average difference over six SAT forms is -.12 raw scores. The effect is also minimal for Hispanic and Asian American examinees, relative to White examinees. However, weighting item responses by difficulty is slightly advantageous for Black examinees relative to White examinees (the average difference over seven forms is .32), which means that Black examinees tend to receive credit for a slightly more difficult set of items, relative to similarly performing White examinees on SAT mathematical tests.

GMAT Verbal Tests. The data shown in Table 16 indicate that weighting of item responses by difficulty has an inconsistent effect on differences in performance. The effects for each focal group and reference group comparison are form specific so it is not possible to interpret the findings.

GMAT Quantitative Tests. The data shown in Table 17 indicate that weighting of item responses by difficulty has little effect on differences in GMAT quantitative scores for males and females; the average difference over three GMAT forms is -.11 raw scores. The effect is also be negligible for Hispanic examinees, relative to White examinees. However, weighting item responses by difficulty is advantageous for Black examinees relative to White examinees, and Asian American examinees relative to White examinees. These results indicate that Black examinees and Asian American examinees tend to receive credit for a

more difficult set of items, relative to similarly performing White examinees on GMAT Quantitative tests.

Summary of Findings and Implications

The purpose of this study was to assess the disadvantage, if any, of reaching difficult items that appear earlier in a test, at the expense of not reaching easier items that may appear later in a test. The study focused on SAT and GMAT data and examined the effect for the following focal groups: females, Asian American, Blacks, and Hispanics.

As expected, the impact of differential test speededness on subgroup differences in test scores is minimal for SAT Mathematical tests and GMAT Quantitative tests. While the tests are differentially speeded for various focal and reference groups, the items in the sections of these tests are generally ordered from easiest items to hardest items. Consequently, whether or not examinees from different subgroups fail to reach easier items at the end of test section is not a factor for these tests.

Also as expected, there appears to be a relationship between the presence of differential speededness on the SAT Verbal test and subgroup differences in test performance. Comparative data on test speededness indicates that the SAT Verbal test is more speeded for Black and Hispanic examinees, relative to White examinees of comparable ability. This observation is consistent with findings of Dorans, Schmitt and Bleistein (1992). A consistent effect of both differential not reaching and differential omitting was observed for the SAT Verbal tests. The findings were replicated over several test forms. Specifically, Black and Hispanic examinees appear to be receiving credit for more difficult items in the SAT Verbal test, relative to White examinees. For Hispanic examinees, this is partly a function of differentially higher rates of not reaching items at the end of

the test sections, relative to White examinees receiving the same formula score. For Black examinees, this is a function of two factors: differentially higher rates of not reaching items at the end of test sections **and** differentially higher rates of omitting items.

Because the SAT Verbal test is composed of four different item types, some of the items toward the end of the test section are relatively easy items. When item responses are weighted by item difficulty, rather than by unit weights, Hispanic and Black examinees tend to perform better than White examinees of comparable ability (matched with respect to formula score on the test). This finding implies that the SAT Verbal test underestimates the formula score that would be obtained by Black and Hispanic test takers had these groups responded to more of the easier items in the test sections. As noted, Black examinees also tend to omit items at a higher rate than White examinees with comparable scores on SAT verbal tests, and this may partially explain why Black examinees receive credit for a more difficult set of items than White examinees. However, this could not explain why Hispanic examinees tend to receive credit for a more difficult set of items than similarly scoring White examinees, because Hispanic examinees tend to omit at a lower rate than comparable White examinees.

While SAT Verbal tests are differentially more speeded for Black and Hispanic examinees, the tests are comparably speeded, and in some cases slightly less speeded, for Asian American examinees, after taking differences in ability into account. Nevertheless, Asian American examinees also receive credit for a more difficult set of items compared to White examinees with the same formula score on the test. A follow-up analysis indicates that Asian American examinees tend to omit at a higher rate than comparably able White examinees; this finding is consistent with results of Schmitt, Dorans, Crone and Maneckshana (1991)

which looked at differential speededness and omit patterns on the SAT. Moreover, the higher rate of omitting is associated with items covering the full range of item difficulty, including some fairly easy items. Thus, while Asian American examinees also receive credit for a more difficult set of items than comparably able White examinees, the effect is primarily associated with differential rates of omitting rather than differential rates of not reaching easier items at the end of the test sections.

Findings for the GMAT Verbal tests are not as clear-cut as the findings for the SAT Verbal tests. On the basis of the three GMAT forms examined in this study, the Verbal test appears to be more speeded for Black examinees relative to comparably able White examinees. Because several items at the end of two of the sections (Reading Comprehension and Sentence Correction) are relatively easy items, an effect similar to the effect for the SAT verbal section was expected. This was not the case and results did not agree across forms. Considerably smaller sample sizes for Black examinees on the GMAT Verbal test relative to the SAT Verbal test may account for the inconsistent results.

In their study of differential test speededness, Dorans et al (1992, p. 316) make the following statements:

> "Differential speededness is an undesirable test property. Because differential speededness can contribute to the incidence of content-free DIF on the keyed option, its impact on test scores needs to be investigated more fully."

The finding from this study -- that Black and Hispanic examinees tend to receive credit for a more difficult set of items than White examinees with similar formula scores on SAT-Verbal -- is particularly related to that issue.

This research attempted to provide a technique for assessing the circumstances under which the effect of test speededness has a differential effect

on the performance of subgroups of a population, when differences in ability are taken into consideration. As is the case with all non-experimental research on test speededness, the results of this study cannot predict how examinees would have performed with a shorter test, or more liberal time limits, because test strategies might have been different under those conditions. Nevertheless, one implication of the findings for the SAT Verbal test relates to the extent information needs to be shared with candidates. Perhaps candidates need to receive clearer information about the manner in which items are ordered within test sections along with advice to attempt the easier items within each item set.

As testing programs move toward increased use of item sets, or testlets, the effect of differential test speededness on test scores will persist. However, it may also be the case that another innovation, increased use of computerized adaptive testing, could have a positive effect for subgroups that are currently disadvantaged by differential test speededness. This is because the phenomenon of not reaching items at the end of a test section should not occur on a computerized adaptive test.

References

Brennan, R. L. (1992). The context of context effects. *Applied Measurement in Education, 5*, 225-264.

Cronbach, L. J. (1984). *Essentials of psychological testing*. New York: Harper.

Dorans, N. J., Schmitt, A. P., & Bleistein, C. A. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement, 29*, 309-319.

Dorans, N. J., Schmitt, A. P., & Curley, W. E. (1988, April). *Differential speededness: Some items have DIF because of where they are, not of what they are.* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

Leary, L. F. & Dorans, N. J. (1985). Implications for altering the context in which test items appear: A historical perspective on an immediate concern. *Review of Educational Research, 55*, 387-413.

Schmitt, A. P. & Bleistein, C. A. (1987). *Factors affecting differential item functioning for Black examinees on Scholastic Aptitude Test analogy items.* (RR-87-23). Princeton, NJ: Educational Testing Service.

Schmitt, A. P., Dorans, N. J., Crone, C. R., & Maneckshana, B. T. (1991) *Differential speededness and item omit patterns on the SAT.* (RR-91-50). Princeton, NJ: Educational Testing Service.

Wang, M. W. & Stanley, J. C. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research, 40*, 663-705.

Table 1

Sample Sizes by SAT and GMAT Test Form and Subgroup

SAT	Admin	Females	Males	Asian	Black	Hispanic	White
A	10/90	56,305	42,120	6,754	9,020	5,879	74,310
B	10/90	30,451	23,023	4,165	5,473	3,834	38,545
C[5]	11/90	96,412	77,215	7,410	16,173	6,899	138,712
D	12/90	36,443	31,885	3,912	9,638	4,050	48,696
E	12/90	35,164	31,232	3,705	9,427	3,875	47,400
F	3/91	81,833	72,868	7,649	17,184	5,752	120,345
G	5/91	150,634	130,575	14,014	20,791	11,043	227,841
GMAT	**Admin**	**Females**	**Males**	**Asian**	**Black**	**Hispanic**	**White**
X	6/91	10,303	14,684	742	1,774	361	21,658
Y	10/91	11,141	15,287	868	2,030	349	22,724
Z	1/92	5,367	8,021	433	966	N/A[6]	11,567

[5]Analyses of SAT Form C were performed for the SAT Verbal test only.

[6]Fewer than 200 Hispanic examinees took GMAT Form Z, so data for this group are not reported.

BEST COPY AVAILABLE

Table 2a

Percent Correct by Item Position: SAT Verbal Section 1

Item	Form A	Form B	Form C	Form D	Form E	Form F	Form G
1	90	94	95	91	83	90	96
2	92	87	86	85	87	84	92
3	88	84	84	73	71	86	87
4	79	76	77	72	68	70	77
5	75	68	80	65	67	64	79
6	71	58	63	72	74	65	77
7	66	56	60	53	48	38	67
8	64	34	38	35	36	66	49
9	60	28	40	24	32	49	21
10	51	26	28	28	29	42	26
11	29	22	17	23	32	36	31
12	28	30	24	25	20	32	24
13	24	28	34	24	25	27	17
14	22	22	19	18	20	24	13
15	9	14	13	12	12	12	10
16	90	93	90	91	87	94	93
17	88	85	89	80	80	86	83
18	75	68	85	81	72	86	71
19	81	63	81	51	82	82	76
20	61	43	79	37	75	85	71
21	40	54	86	39	46	77	54
22	25	34	58	37	21	63	34
23	19	37	36	14	18	29	36
24	16	33	21	18	26	24	34
25	20	39	19	8	16	14	22
26	68	74	83	79	49	50	58
27	80	76	85	77	59	40	57
28	81	65	58	55	56	25	55
29	75	39	70	59	57	38	71
30	61	44	35	77	53	39	66
31	46	64	36	48	56	33	68
32	30	53	34	40	39	32	64
33	47	52	52	26	25	41	42
34	40	54	33	51	24	24	67
35	42	40	37	37	23	43	46
36	97	93	91	84	84	95	95
37	83	84	81	84	84	89	94
38	61	76	79	75	70	86	85
39	75	77	65	72	65	82	69
40	57	73	54	70	36	74	41
41	40	39	42	34	34	51	20
42	24	45	39	38	34	44	27
43	35	35	33	24	26	23	37
44	30	37	40	33	33	39	24
45	37	20	37	30	37	46	25

Table 2b

Percent Correct by Item Position: SAT Verbal Section 2

Item	Form A	Form B	Form C	Form D	Form E	Form F	Form G
1	85	92	91	82	90	90	94
2	83	88	83	86	86	90	81
3	75	84	71	80	91	58	74
4	68	61	76	69	70	71	74
5	65	53	58	34	58	44	44
6	30	31	35	25	55	31	49
7	37	29	27	25	30	26	30
8	20	27	27	21	28	17	29
9	28	20	16	20	24	25	19
10	19	19	6	9	11	18	13
11	92	78	91	92	87	85	90
12	72	78	70	73	47	69	67
13	39	72	81	47	23	69	85
14	25	29	28	37	19	57	24
15	21	9	24	27	13	14	20
16	85	90	87	85	90	88	83
17	91	86	82	82	76	78	83
18	81	77	77	74	77	62	84
19	74	76	74	63	72	48	73
20	46	65	70	49	30	37	80
21	36	52	33	31	21	31	67
22	30	64	38	30	22	29	45
23	26	24	30	32	24	18	37
24	18	15	22	22	15	28	43
25	20	22	20	14	26	17	16
26	84	52	93	81	48	72	62
27	95	64	54	77	58	68	87
28	91	50	86	66	66	86	56
29	76	28	78	65	47	56	87
30	29	75	58	53	71	52	37
31	75	59	33	35	49	69	79
32	55	23	55	73	62	80	38
33	50	63	74	45	65	46	45
34	28	37	67	52	71	53	62
35	47	76	25	36	43	38	53
36	62	79	77	31	72	25	40
37	40	34	58	53	68	46	64
38	20	76	36	48	41	66	67
39	53	85	28	28	34	70	58
40	27	48	19	47	49	37	37

Table 3a

Percent Correct by Item Position: SAT Mathematical Section 1

Item	Form A	Form B	Form D	Form E	Form F	Form G
1	95	88	90	89	93	79
2	96	97	79	69	85	95
3	83	75	89	66	78	81
4	71	72	76	74	76	83
5	85	74	76	68	84	80
6	71	77	64	66	66	93
7	71	78	69	70	84	75
8	76	82	73	61	80	80
9	77	66	70	66	76	84
10	72	69	68	62	55	84
11	59	69	76	54	68	70
12	45	71	47	68	70	67
13	70	59	49	83	65	61
14	58	62	37	84	60	72
15	57	49	72	56	61	49
16	45	34	42	53	45	37
17	47	46	65	45	16	43
18	33	32	35	32	50	41
19	42	34	25	29	56	50
20	34	79	20	36	34	21
21	28	69	22	25	35	36
22	12	17	14	27	28	41
23	27	26	18	18	24	17
24	9	31	20	11	24	11
25	22	15	20	13	28	25

Table 3b

Percent Correct by Item Position: SAT Mathematical Section 2

Item	Form A	Form B	Form D	Form E	Form F	Form G
1	98	90	84	90	87	84
2	83	89	91	77	95	95
3	81	75	76	74	79	71
4	84	83	78	72	83	82
5	64	50	67	84	79	64
6	69	41	63	74	74	76
7	63	78	76	67	52	65
8	84	90	90	96	79	94
9	92	85	74	87	70	73
10	77	89	72	61	78	83
11	86	82	49	74	75	77
12	64	79	55	67	81	71
13	70	49	45	67	61	84
14	76	62	49	55	66	75
15	62	77	49	63	61	69
16	62	68	62	61	60	66
17	68	68	55	46	49	73
18	71	67	53	55	37	63
19	65	57	50	40	42	44
20	46	43	45	30	74	58
21	49	60	33	59	53	55
22	59	66	47	40	25	52
23	24	25	42	31	45	22
24	30	37	32	38	43	34
25	27	27	24	17	33	57
26	31	20	29	16	27	27
27	24	30	24	22	25	29
28	66	86	49	57	64	33
29	54	66	45	38	40	51
30	46	33	53	28	89	53
31	36	46	21	26	21	38
32	30	16	33	20	26	14
33	20	26	19	10	27	23
34	16	21	18	15	20	29
35	33	11	14	23	13	18

Table 4

Percent Correct by Item Position: GMAT Verbal Sections

Sect.	Reading Comp.			Critical Reasoning			Sentence Corr.		
Item	Form X	Form Y	Form Z	Form X	Form Y	Form Z	Form X	Form Y	Form Z
1	66	48	89	94	91	89	97	83	84
2	66	86	89	59	87	89	82	85	80
3	78	59	81	88	93	93	84	84	88
4	77	49	88	93	89	89	94	83	85
5	51	69	73	89	90	84	77	84	88
6	54	82	49	75	88	84	82	79	73
7	46	40	41	70	89	79	60	72	74
8	65	76	90	60	78	84	36	69	78
9	66	59	84	71	78	75	42	69	73
10	30	61	87	83	74	71	69	61	78
11	58	78	87	65	60	74	76	73	76
12	41	53	49	51	71	55	24	59	71
13	56	61	36	64	77	59	58	50	65
14	40	42	29	60	63	62	45	34	65
15	69	58	30	61	55	57	69	46	58
16	43	62	59	59	46	60	83	25	59
17	50	71	73	35	36	45	44	20	38
18	77	88	75	50	56	48	76	56	59
19	46	67	68	29	50	30	45	97	44
20	73	28	81	25	27	36	34	38	38
21	91	73	68				26	38	51
22	77	59	59				57	72	32
23	28	19	48				46	51	68
24	25						70	41	33
25	92						47	30	63
26								60	56
27								77	63

Table 5

Percent Correct by Item Position: GMAT Quantitative Sections

Sect.	Prob. Solving 1			Prob. Solving 2			Data Sufficiency		
Item	Form X	Form Y	Form Z	Form X	Form Y	Form Z	Form X	Form Y	Form Z
1	92	94	88	92	91	97	85	94	98
2	87	90	84	88	83	88	83	87	78
3	88	92	90	91	77	78	77	92	91
4	82	77	85	70	94	78	79	84	86
5	69	85	76	75	73	66	66	60	82
6	59	75	82	68	81	65	79	72	69
7	76	65	62	55	73	68	66	63	80
8	75	56	56	72	52	63	56	74	69
9	46	57	65	52	59	44	54	67	58
10	59	48	56	63	47	70	53	46	62
11	55	57	57	55	34	60	55	44	60
12	50	37	59	40	52	77	47	56	61
13	51	41	46	46	53	52	41	63	63
14	51	58	49	46	52	37	35	39	56
15	53	48	28	39	27	48	56	73	45
16	45	47	22	31	42	37	36	40	48
17	26	32	29	30	21	34	24	29	41
18	24	41	26	21	41	27	35	47	46
19	43	18	47	29	25	26	29	27	41
20	21	35	69	46	31	37	25	20	42
21							37	34	31
22							31	46	21
23							16	22	24
24							16	32	21
25							33	23	20

Table 6

Average Number of Items Not Reached, on 85-Item SAT Verbal Test, by Form and Subgroup

SAT	Admin	Females	Males	Asian	Black	Hispanic	White
A	10/90	2.2	2.3	1.8	4.0	3.2	2.0
B	10/90	2.5	2.7	2.3	4.2	3.3	2.4
C	11/90	2.5	2.5	2.2	4.2	3.5	2.3
D	12/90	2.5	2.6	2.3	3.9	3.4	2.2
E	12/90	3.2	3.4	2.9	4.8	4.0	3.0
F	3/91	2.3	2.4	1.8	4.3	3.2	2.0
G	5/91	2.0	2.1	1.9	3.7	3.0	1.9

Table 7

Average Number of Items Not Reached on 60-Item SAT Mathematical Test, by Form and Subgroup

SAT	Admin	Females	Males	Asian	Black	Hispanic	White
A	10/90	2.7	2.4	2.2	3.3	2.9	2.5
B	10/90	1.9	1.6	1.7	2.5	2.3	1.6
D	12/90	3.1	3.0	2.7	3.8	3.5	2.8
E	12/90	2.0	1.9	1.9	2.6	2.5	1.8
F	3/91	2.5	2.3	2.1	3.4	3.0	2.3
G	5/91	2.6	2.3	2.2	3.5	3.2	2.3

Note: Analyses of Form C were performed for the SAT Verbal test only.

Table 8

Average Number of Items Not Reached on 70-Item GMAT Verbal Test, by Form and Subgroup

GMAT	Admin	Females	Males	Asian	Black	Hispanic	White
X	6/91	3.3	3.2	3.8	5.9	4.7	3.0
Y	11/91	3.6	3.3	3.4	6.2	4.2	3.2
Z	1/92	3.8	3.6	3.7	6.1	N/A	3.4

Table 9

Average Number of Items Not Reached on 65-Item GMAT Quantitative Test, by Form and Subgroup

GMAT	Admin	Females	Males	Asian	Black	Hispanic	White
X	6/91	3.9	3.3	3.4	5.3	4.9	3.4
Y	11/91	4.3	3.6	3.7	5.1	4.6	3.8
Z	1/92	5.1	4.3	3.7	6.4	N/A	4.5

Table 10

Summary of Average Conditional Differences in Not Reaching Items on SAT Verbal Forms

Focal Group: Reference Group:	Female Male	Black White	Hispanic White	Asian White
A	-.144	1.301	.856	-.110
B	-.351	.842	.391	-.049
C	-.128	1.219	.824	.107
D	-.219	1.180	.832	.211
E	-.350	1.029	.578	-.013
F	-.161	1.299	.678	-.017
G	-.123	.994	.676	.134

Table 11

Summary of Average Conditional Differences in Not Reaching Items on SAT Mathematical Forms

Focal Group: / Reference Group:	Female / Male	Black / White	Hispanic / White	Asian / White
A	.026	.742	.341	.016
B	.124	.791	.622	.209
D	- .118	1.077	.707	.052
E	.009	.858	.704	.240
F	- .026	1.141	.630	.153
G	.068	.964	.721	.189

Note: Analyses of Form C were for the SAT Verbal test only.

Table 12

Summary of Average Conditional Differences in Not Reaching Items on GMAT Verbal Forms

Focal Group: Reference Group:	Female Male	Black White	Hispanic White	Asian White
X	- .005	1.162	.995	.527
Y	.136	1.229	.378	.080
Z	.033	.743	N/A	- .067

Table 13

Summary of Average Conditional Differences in Not Reaching Items on GMAT Quantitative Forms

Focal Group:	Female	Black	Hispanic	Asian
Reference Group:	Male	White	White	White
X	-.007	1.212	1.000	.284
Y	.206	.942	.555	.382
Z	-.072	1.087	N/A	-.137

Table 14

Values of AWCDS on SAT Verbal Forms

Focal Group: / Reference Group:	Female / Male	Black / White	Hispanic / White	Asian / White
A	-.291	.374	.362	.618
B	-.129	.225	.499	.674
C	-.047	.429	.282	.576
D	-.017	.636	.463	.730
E	.160	.447	.312	.837
F	.053	.556	.454	.609
G	.111	.443	.298	.686
Average	-.023	.444	.382	.676

Table 15

Values of AWCDS on SAT Mathematical Forms

Focal Group: Reference Group:	Female Male	Black White	Hispanic White	Asian White
A	-.083	.277	.130	.166
B	-.240	.216	.078	.146
D	-.153	.279	.017	.095
E	-.039	.338	.143	.359
F	-.075	.436	.180	.246
G	-.162	.350	.111	.176
Average	-.125	.316	.110	.198

Note: Analyses of Form C were for the SAT Verbal test only.

Table 16

Values of AWCDS on GMAT Verbal Forms

Focal Group: **Reference Group:**	Female Male	Black White	Hispanic White	Asian White
X	.002	.266	.089	.366
Y	-.092	.114	-.164	-.247
Z	-.371	-.117	N/A	.172
Average	-.154	.088	-.038	.097

Table 17

Values of AWCDS on GMAT Quantitative Forms

Focal Group:	Female	Black	Hispanic	Asian
Reference Group:	Male	White	White	White
X	-.157	.250	-.093	.119
Y	-.085	.379	.122	.279
Z	-.096	.388	NA	.369
Average	-.113	.339	.015	.256

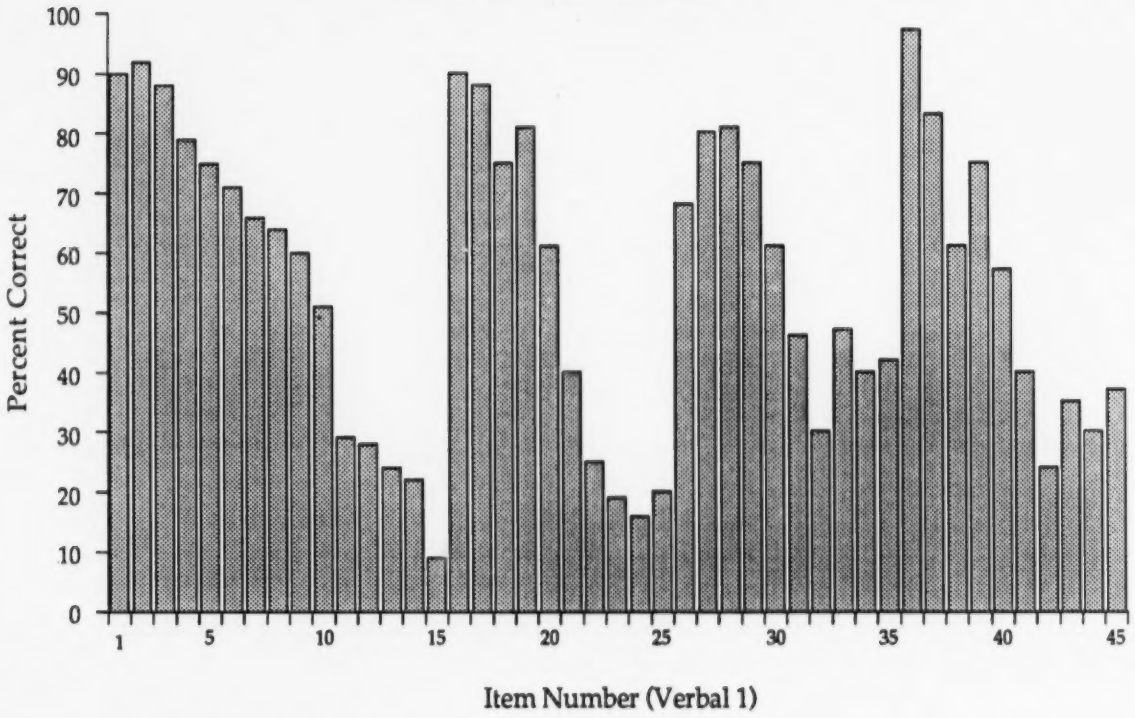
Table 18

Summary of Average Conditional Differences in Omitting Items on SAT Verbal Forms

Focal Group: / Reference Group:	Female / Male	Black / White	Hispanic / White	Asian / White
A	-0.938	0.758	-0.786	0.602
B	-0.982	0.940	-1.205	0.341
C	-0.898	0.804	-0.695	0.919
D	-0.800	1.496	-0.481	1.102
E	-0.821	0.985	-1.128	1.220
F	-0.837	1.310	-0.853	0.758
G	-0.673	0.746	-0.386	0.684

Figure 1

Percent Correct by Item Position
SAT Verbal Form A



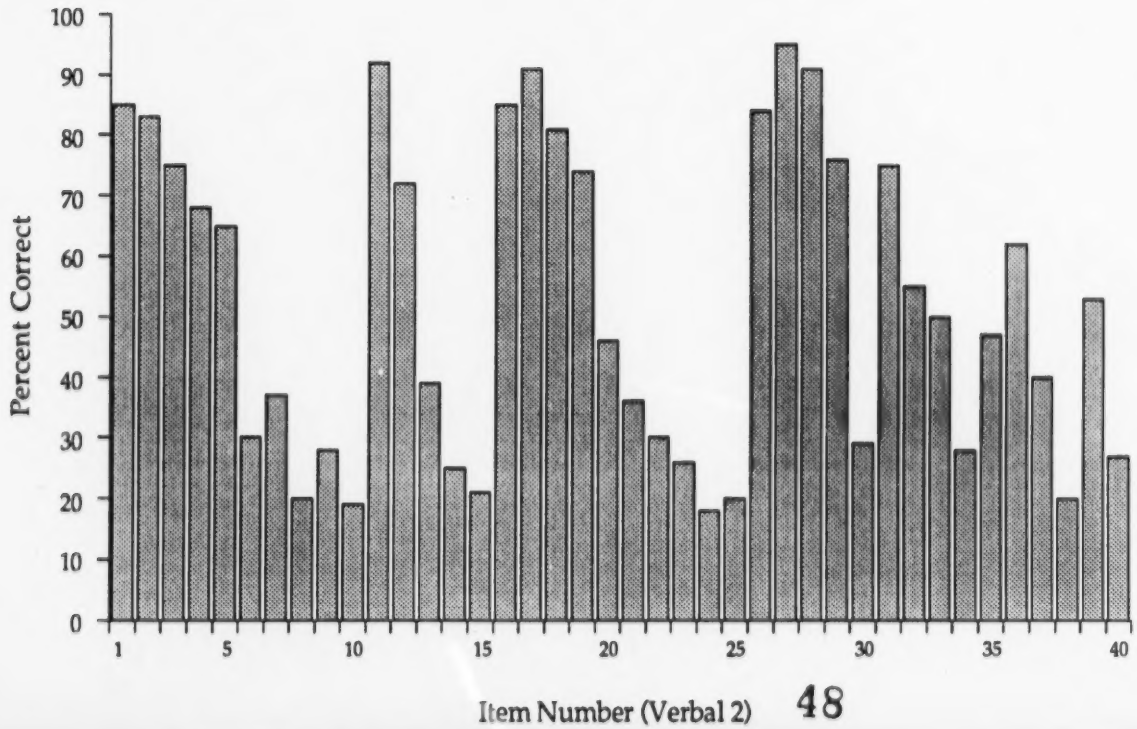
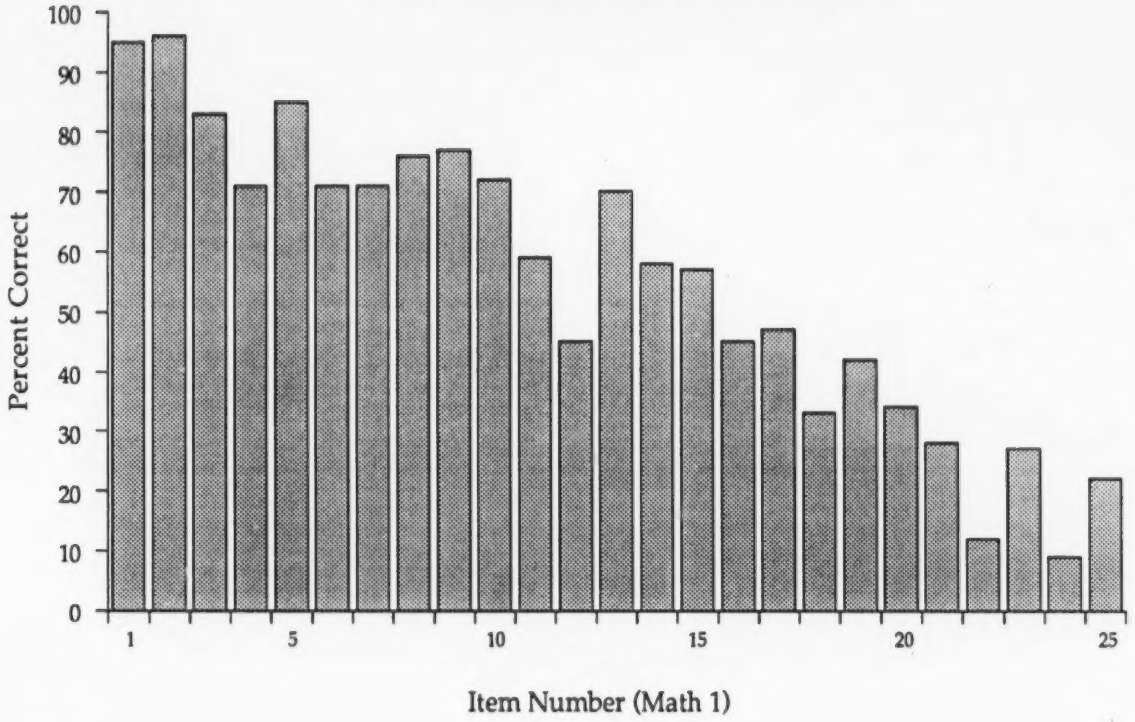


Figure 2

Percent Correct by Item Position
SAT Mathematical Form A



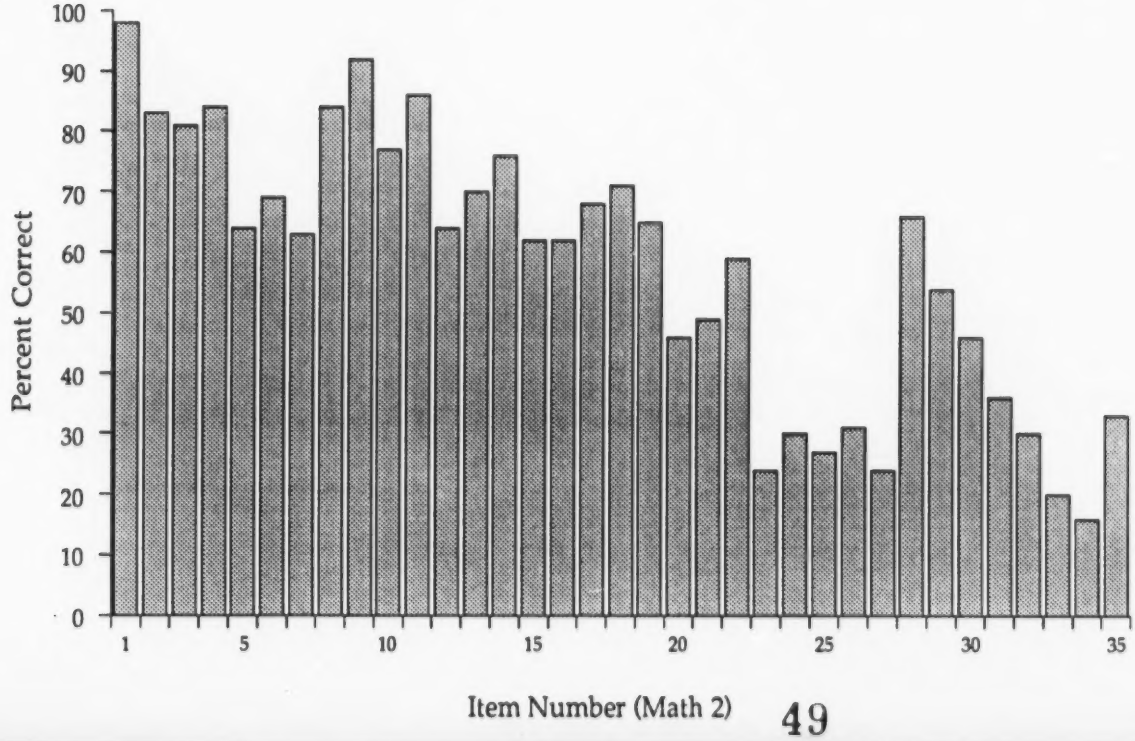
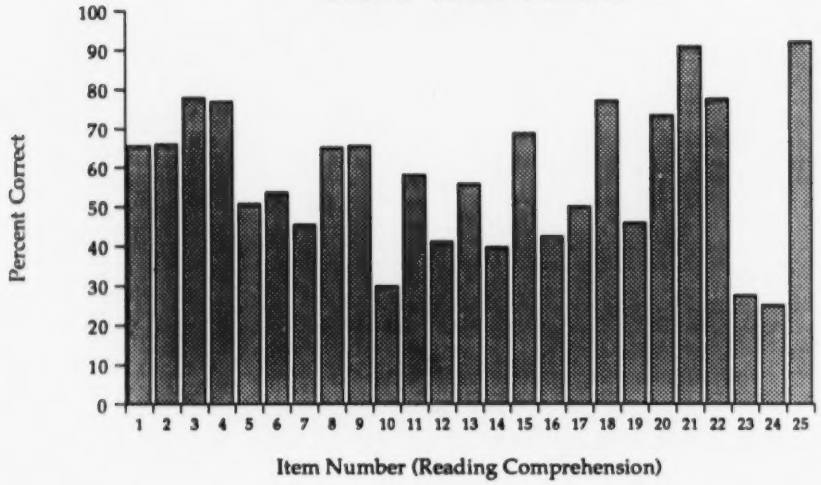
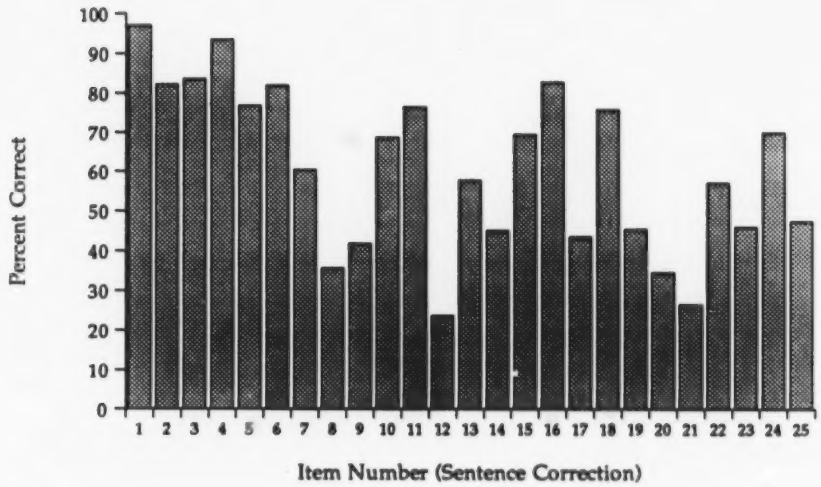


Figure 3

Percent Correct by Item Position
GMAT Verbal Form X





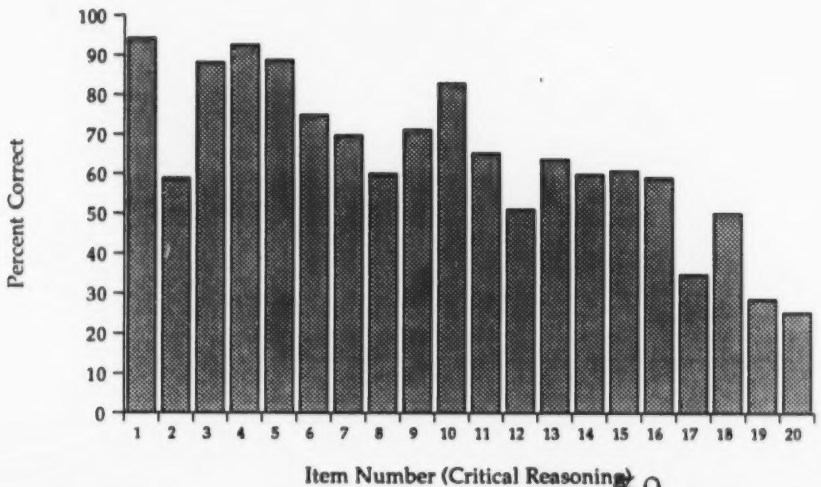
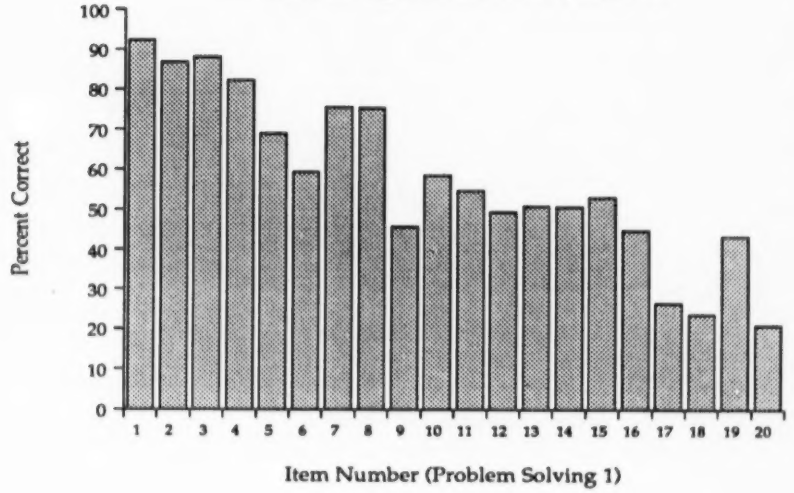
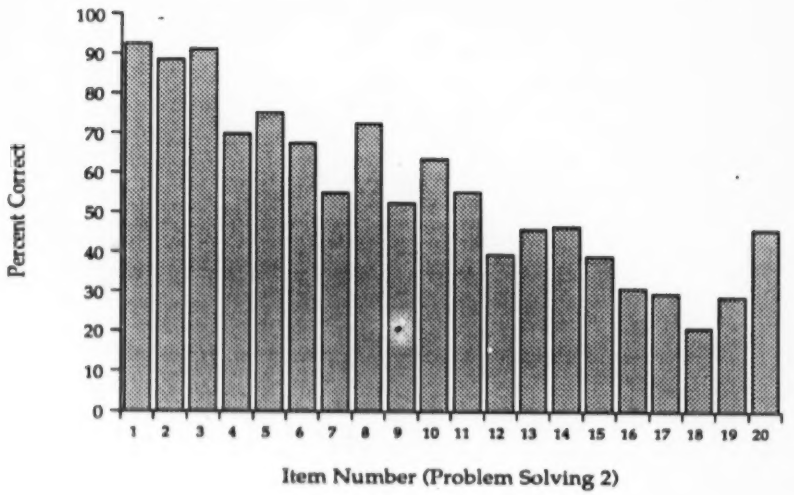


Figure 4

Percent Correct by Item Position
GMAT Quantitative Form X





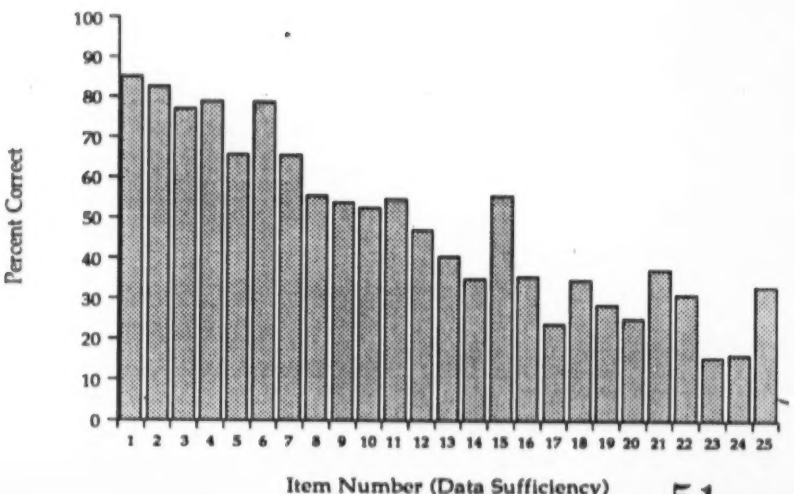


Figure 5: Form A Verbal by Gender

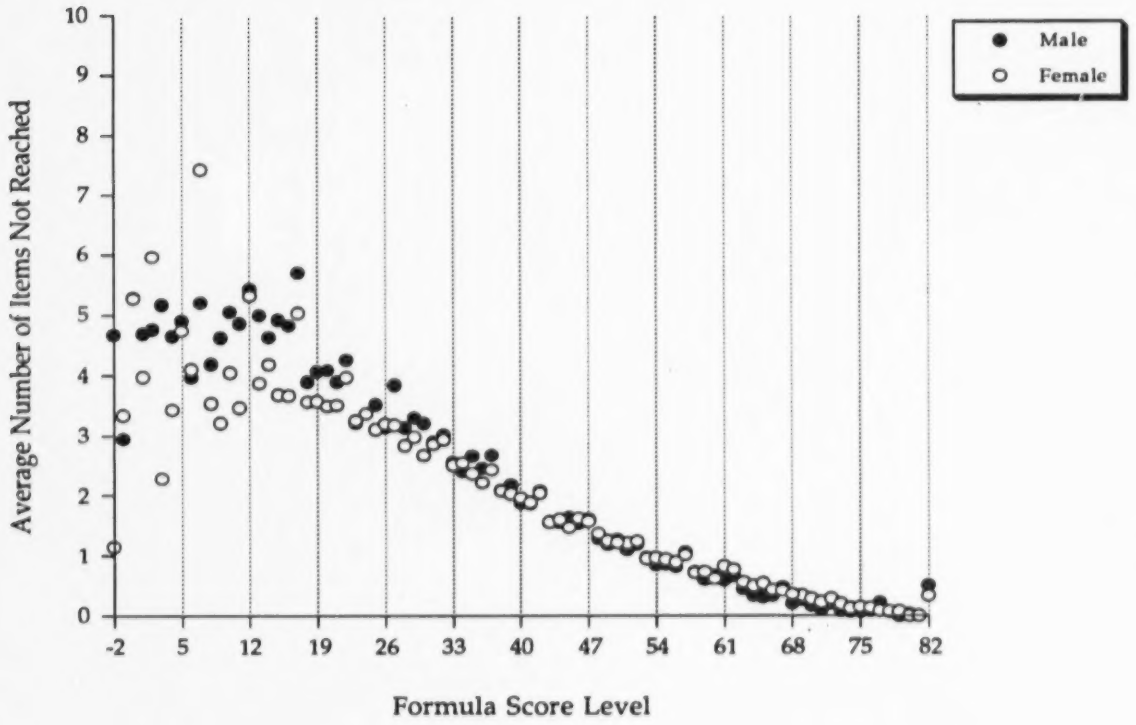


Figure 5: Form A Verbal by Race/Ethnicity

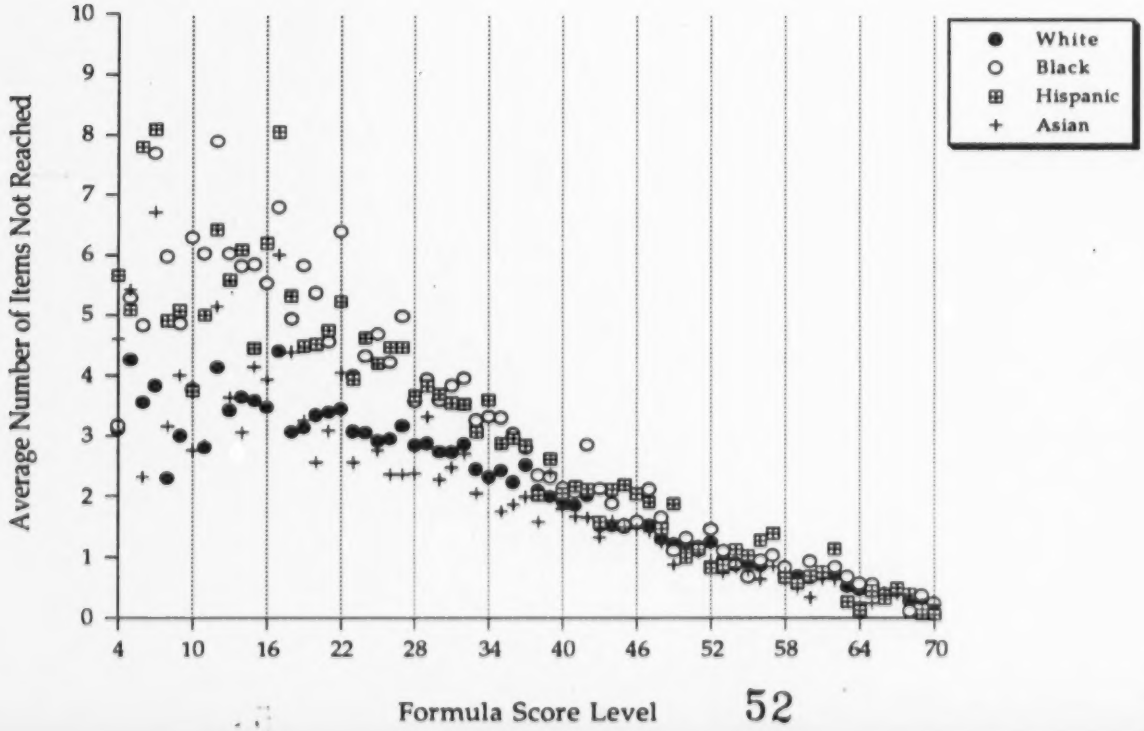


Figure 6: Form B Verbal by Gender

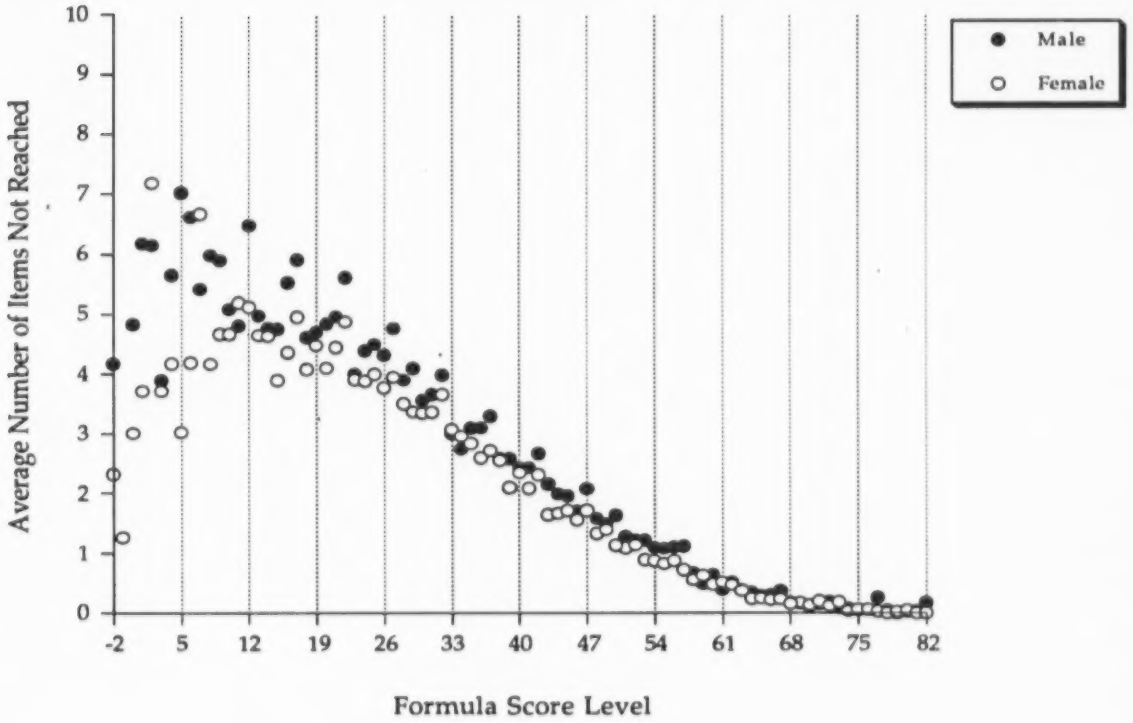


Figure 6: Form B Verbal by Race/Ethnicity

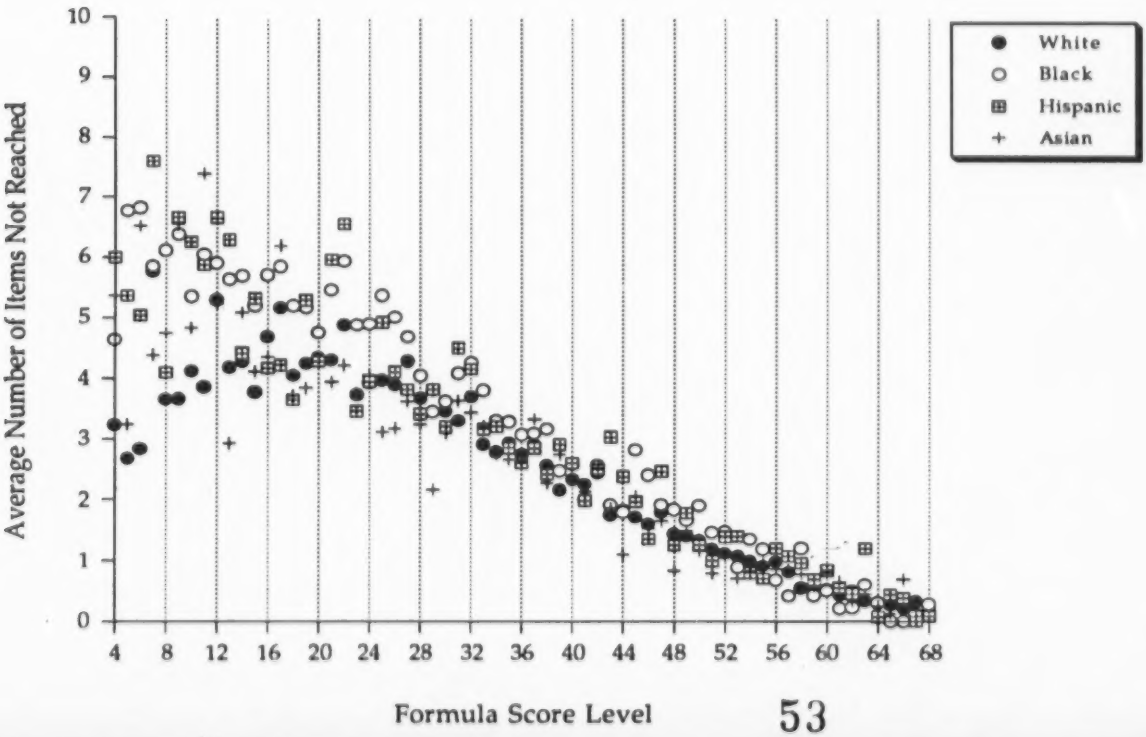


Figure 7: Form C Verbal by Gender

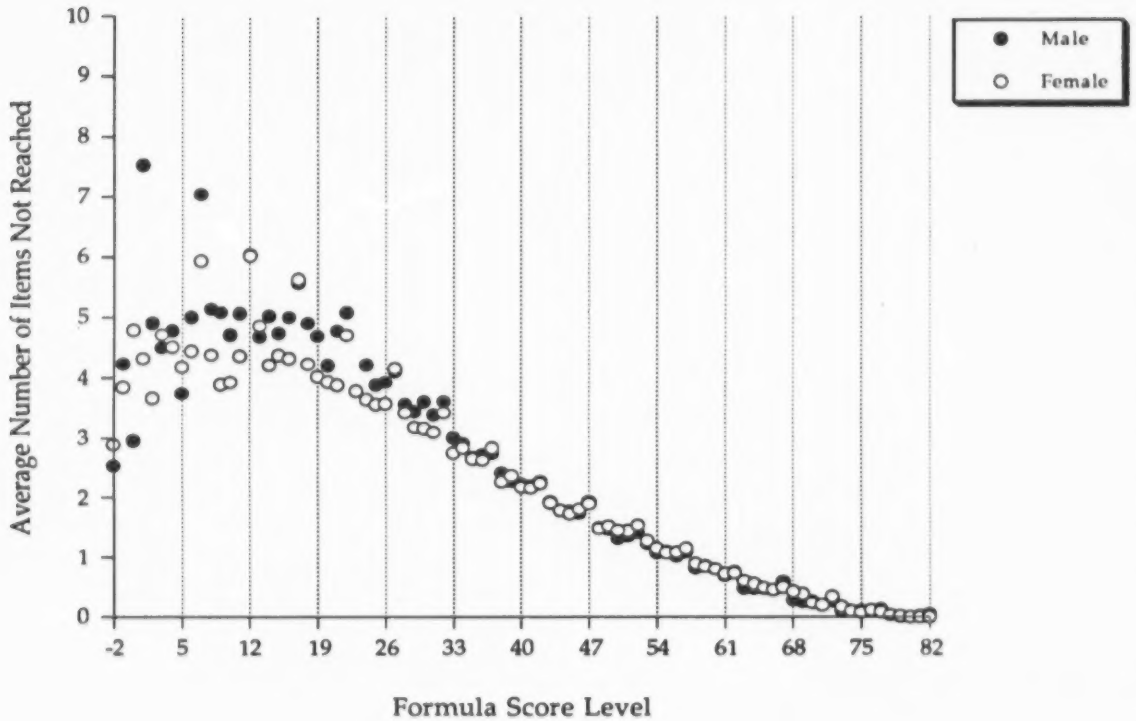


Figure 7: Form C Verbal by Race/Ethnicity

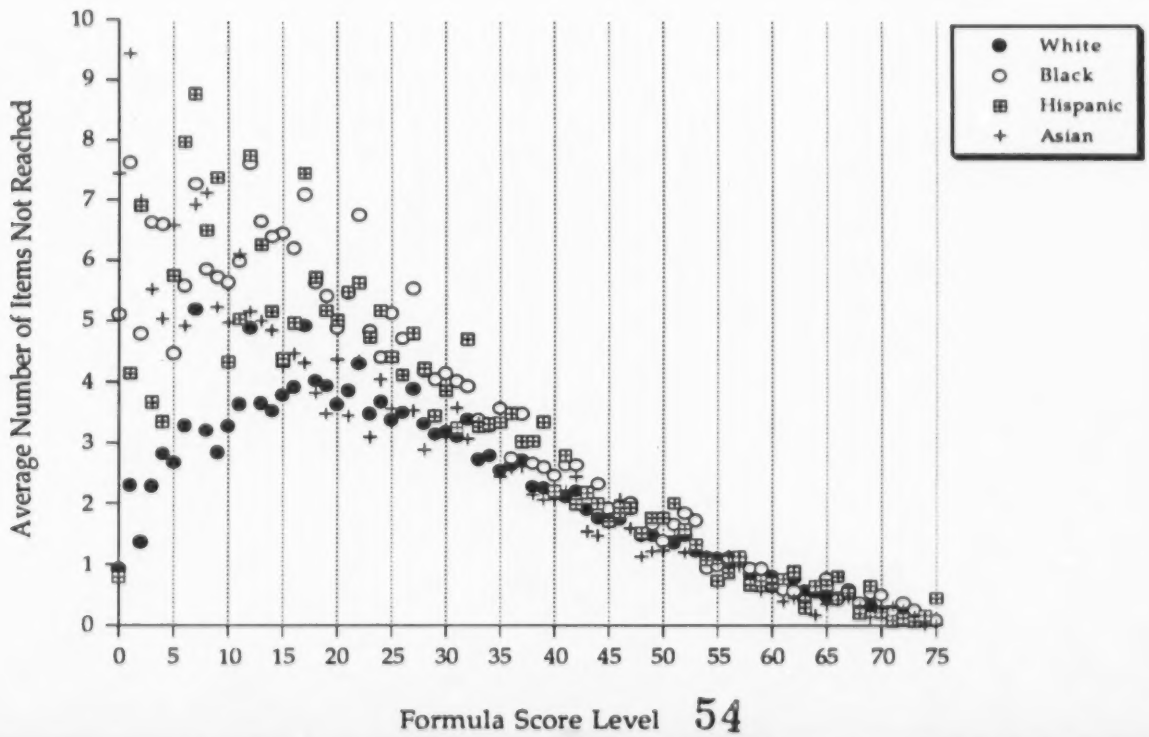


Figure 8: Form D Verbal by Gender

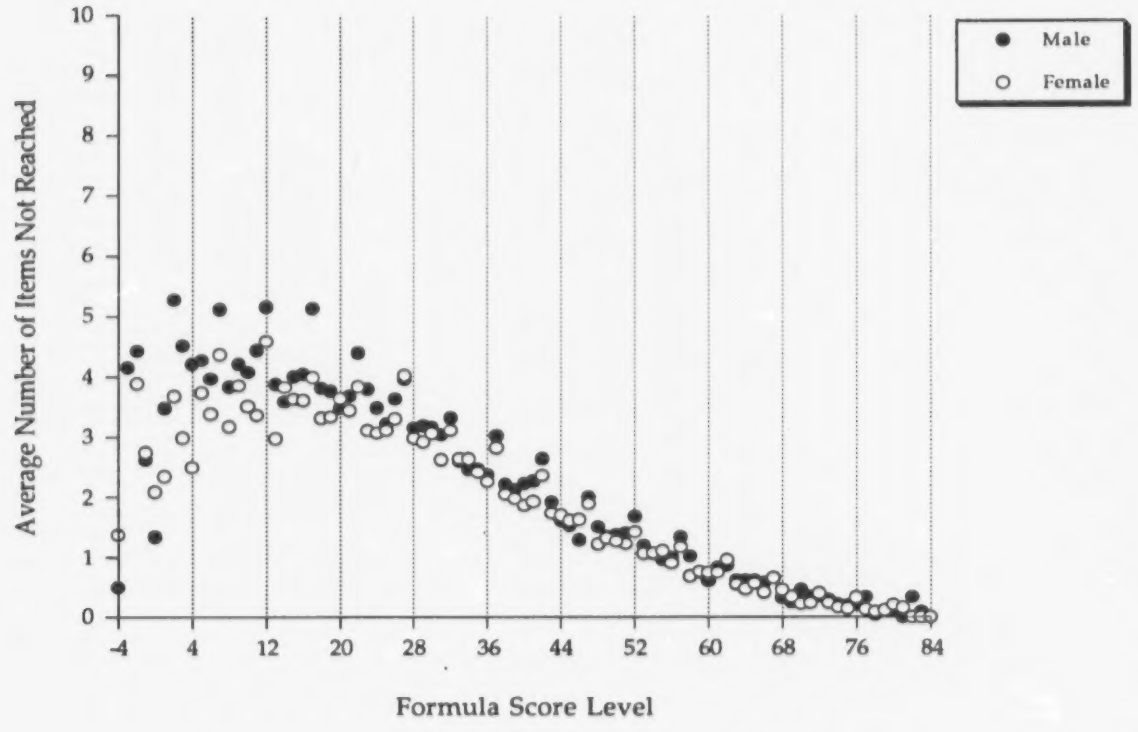


Figure 8: Form D Verbal by Race/Ethnicity

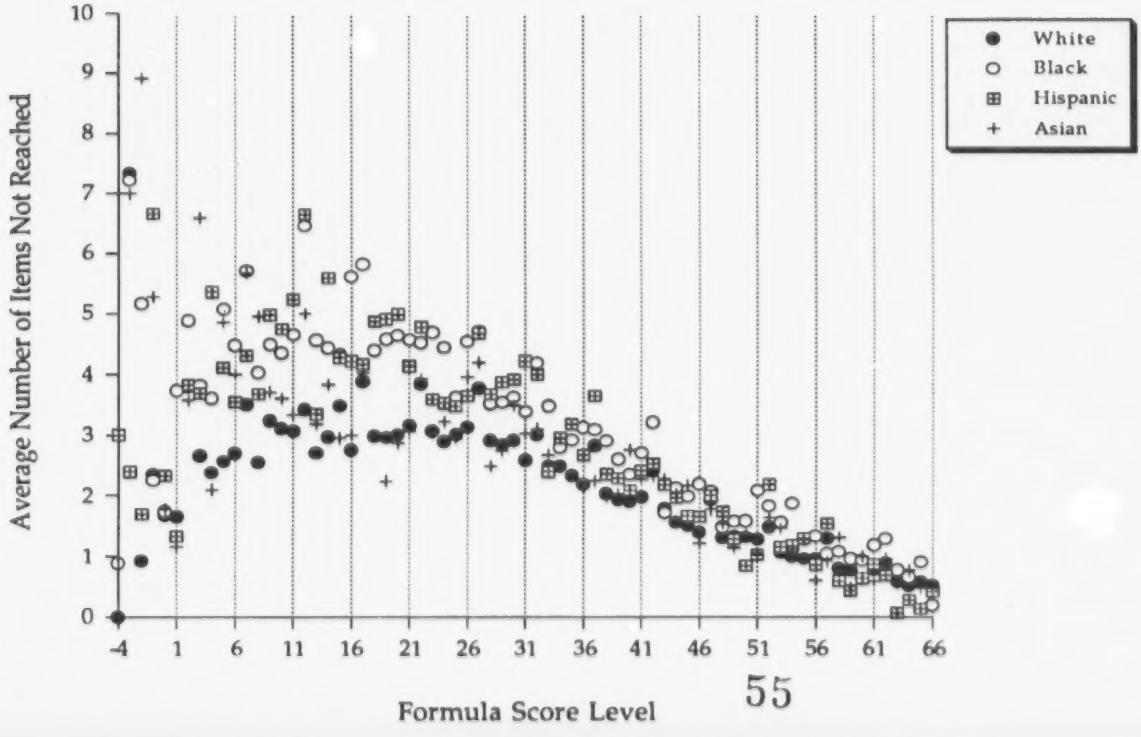


Figure 9: Form E Verbal by Gender

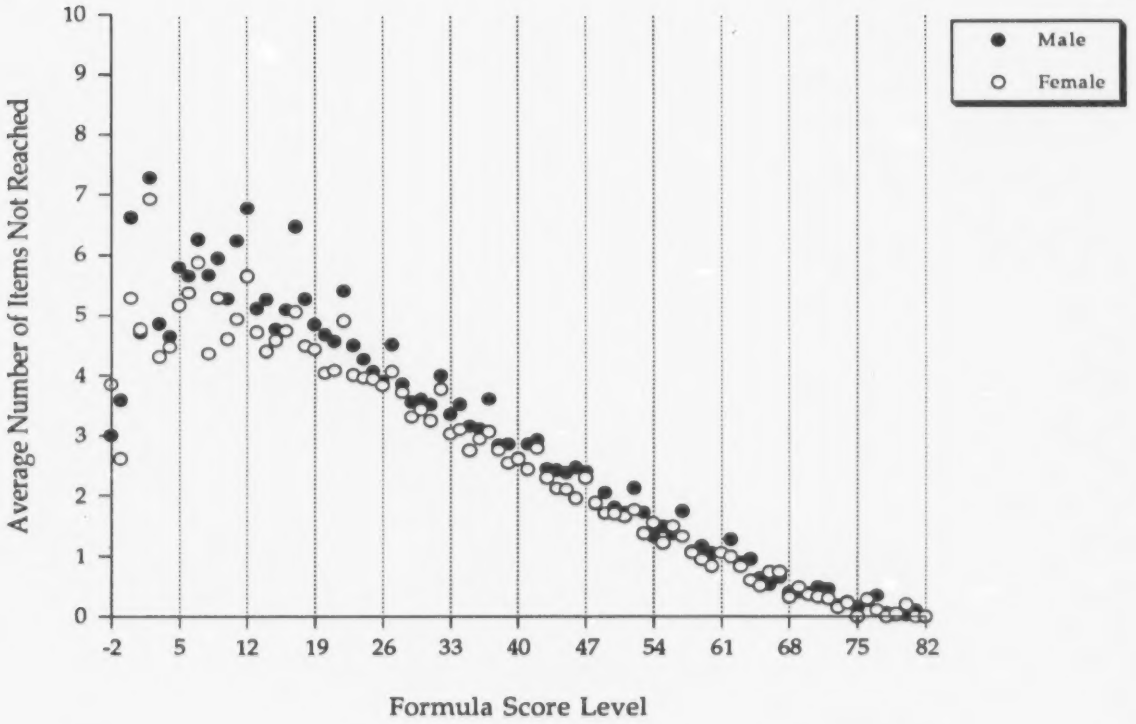


Figure 9: Form E Verbal by Race/Ethnicity

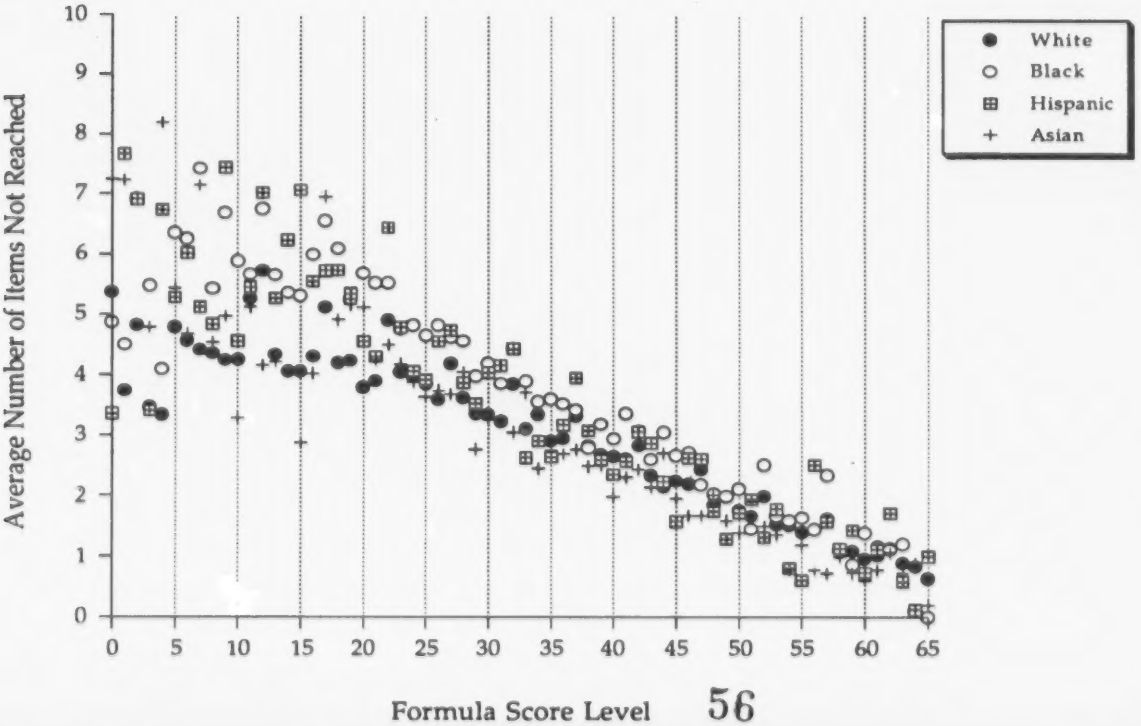


Figure 10: Form F Verbal by Gender

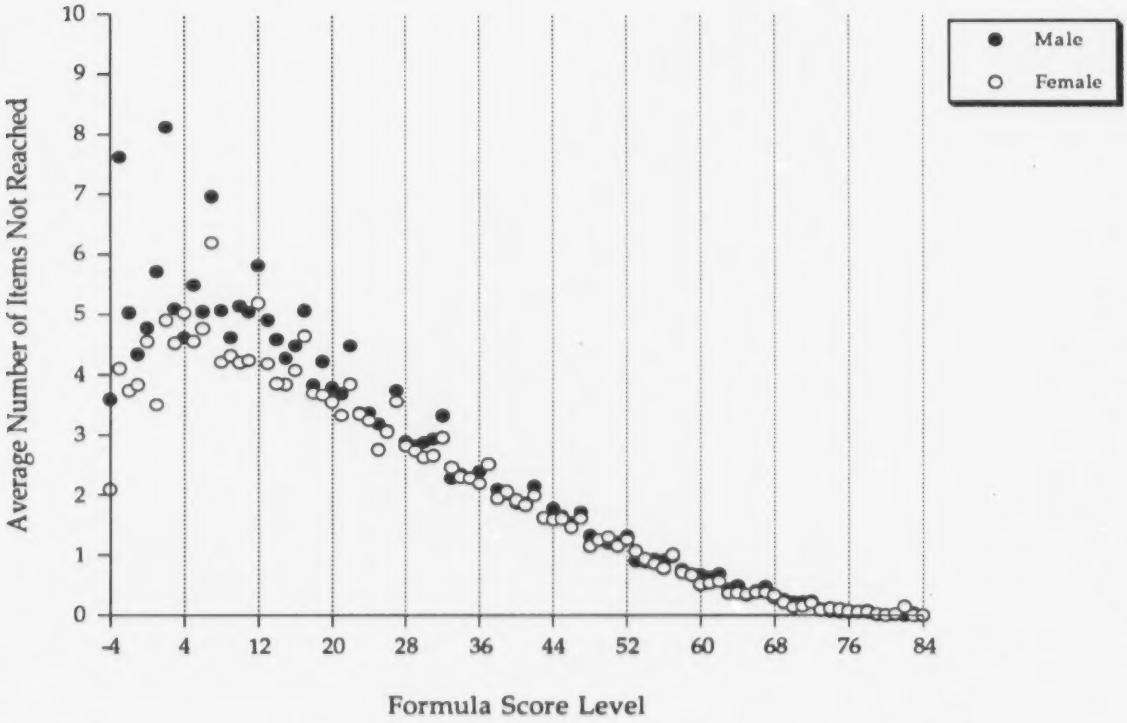


Figure 10: Form F Verbal by Race/Ethnicity

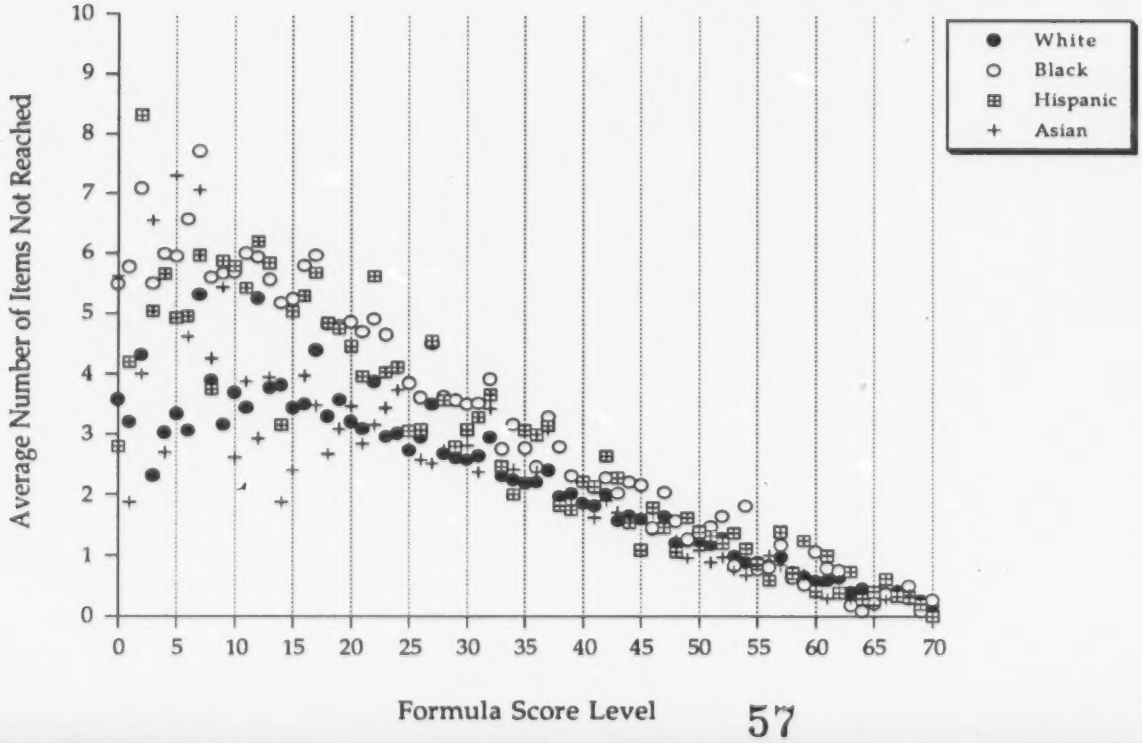


Figure 11: Form G Verbal by Gender

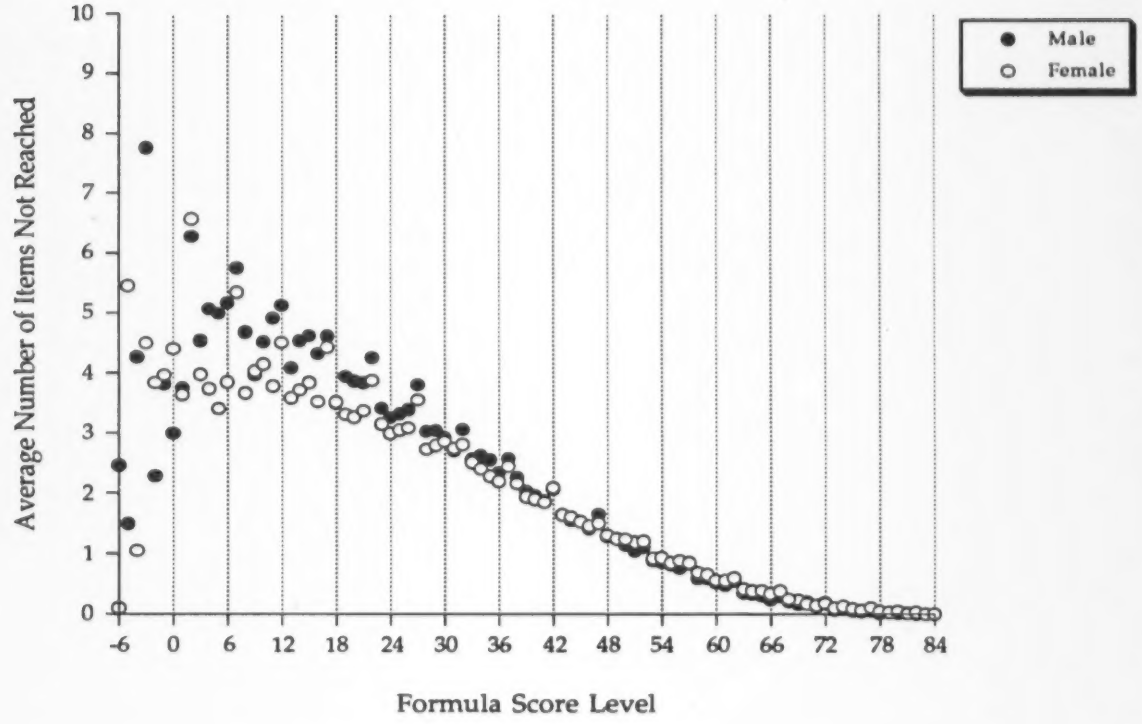


Figure 11: Form G Verbal by Race/Ethnicity

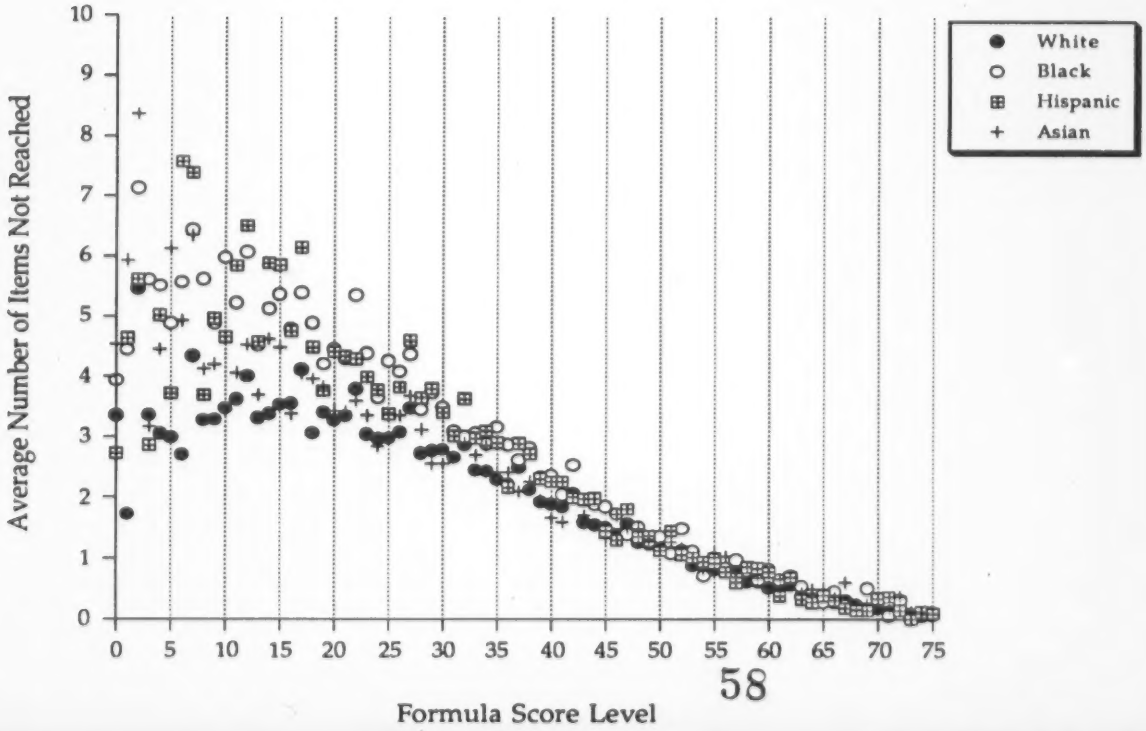


Figure 12: Form A Math by Gender

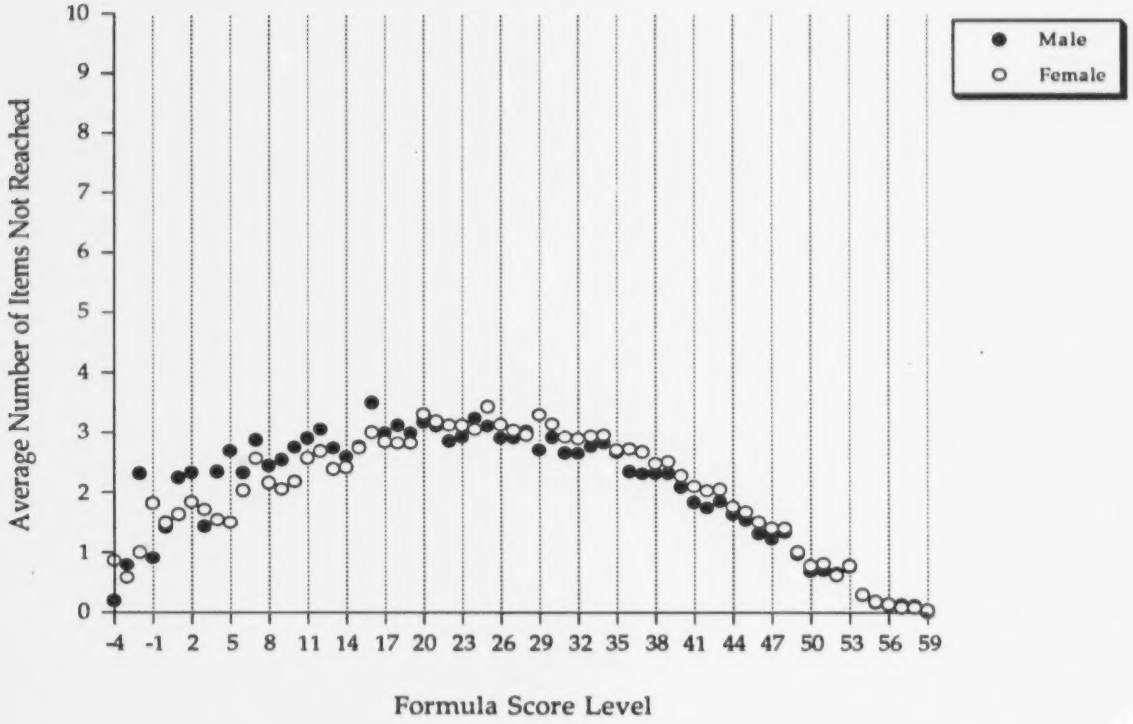


Figure 12: Form A Math by Race/Ethnicity

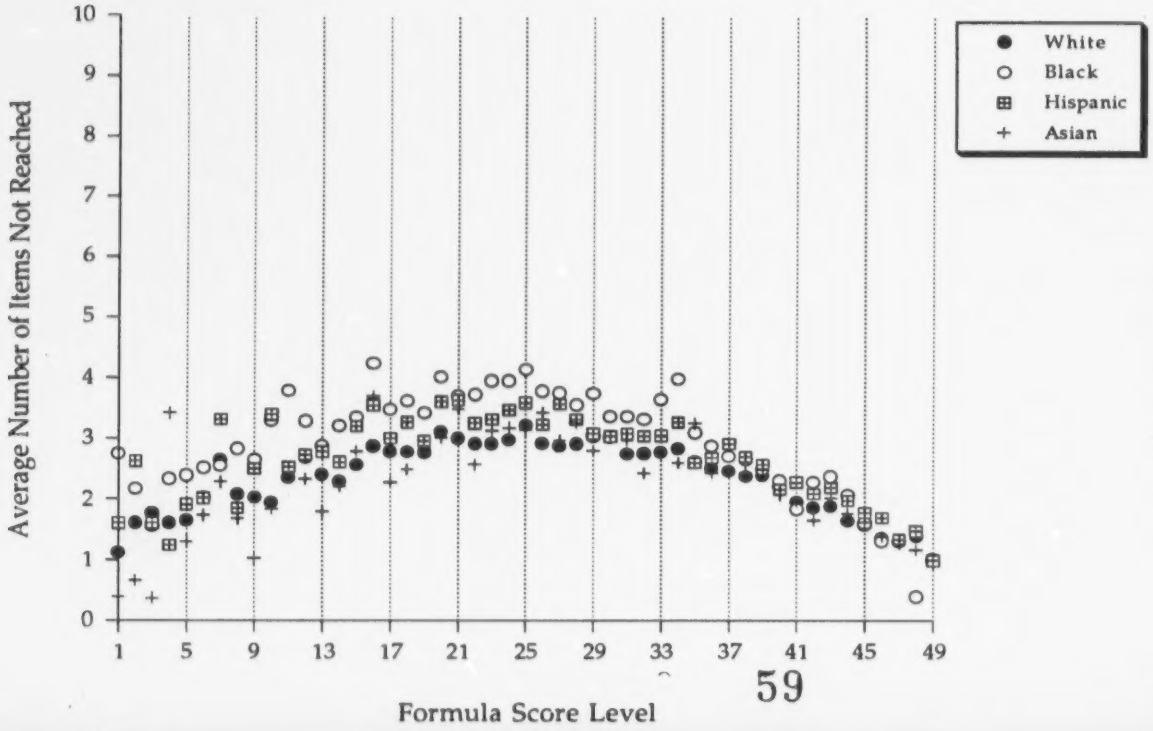


Figure 13: Form B Math by Gender

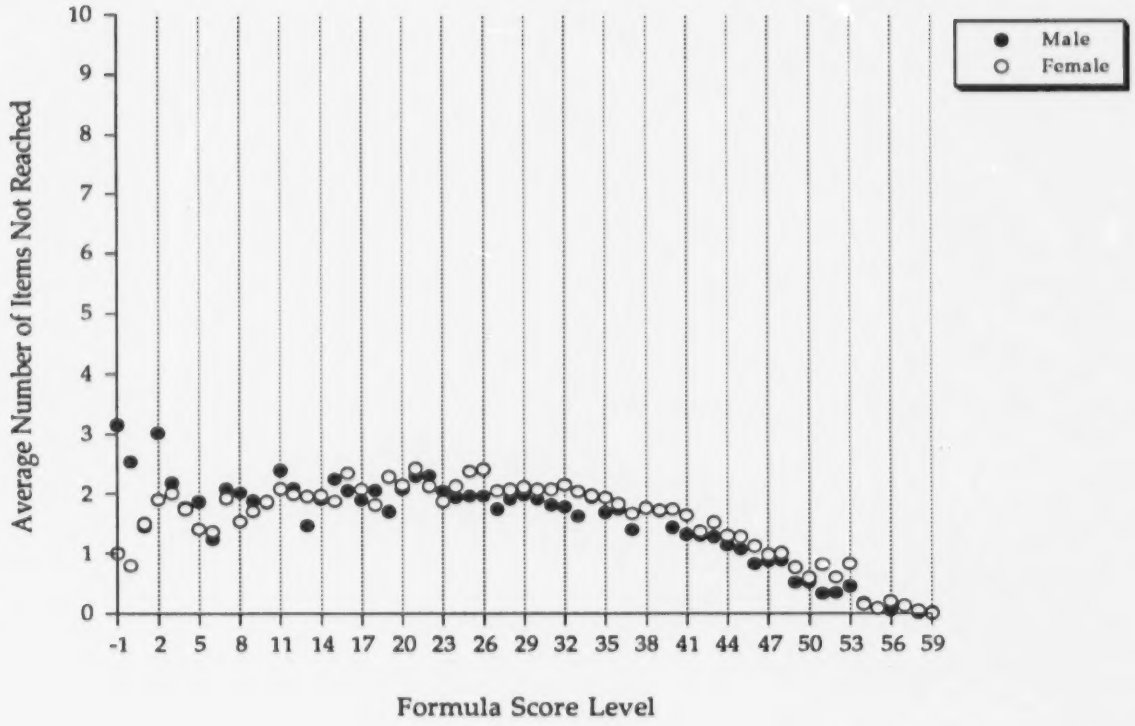


Figure 13: Form B Math by Race/Ethnicity

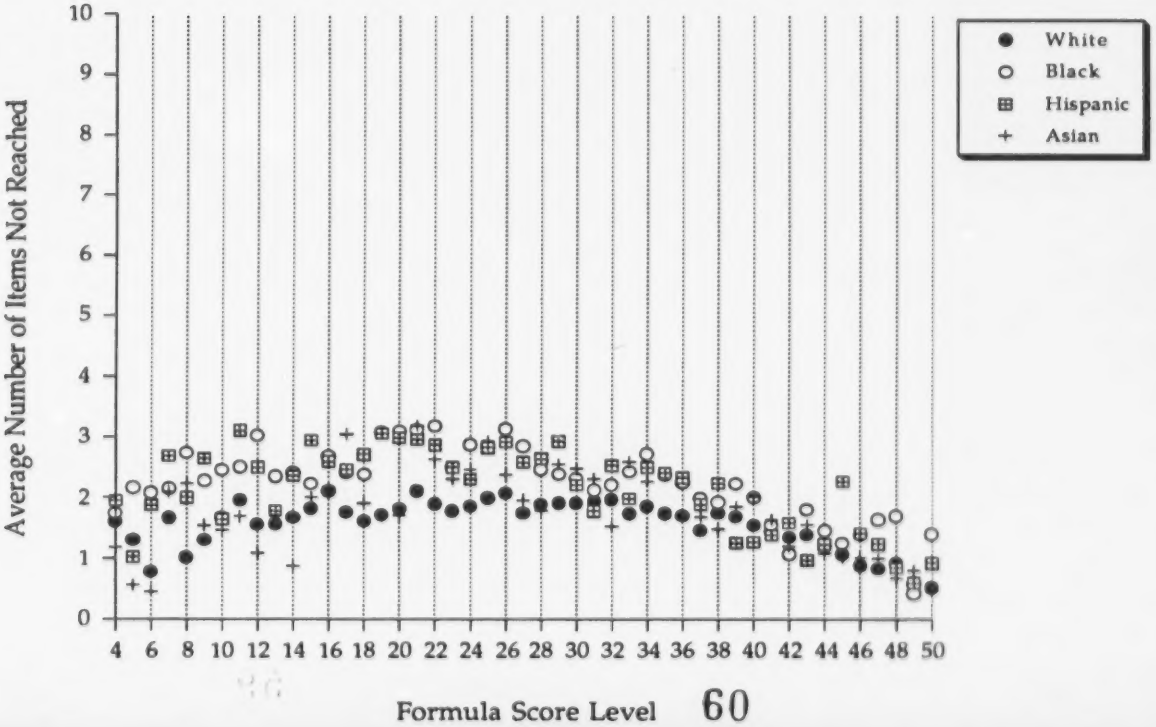


Figure 14: Form D Math by Gender

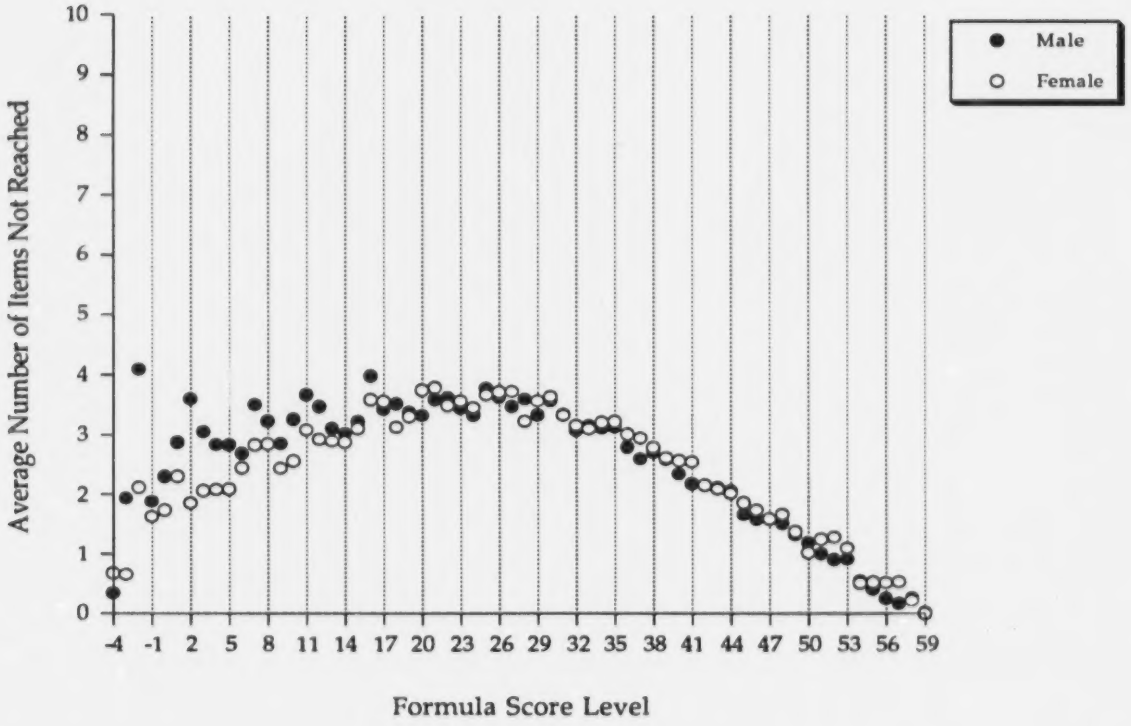


Figure 14: Form D Math by Race/Ethnicity

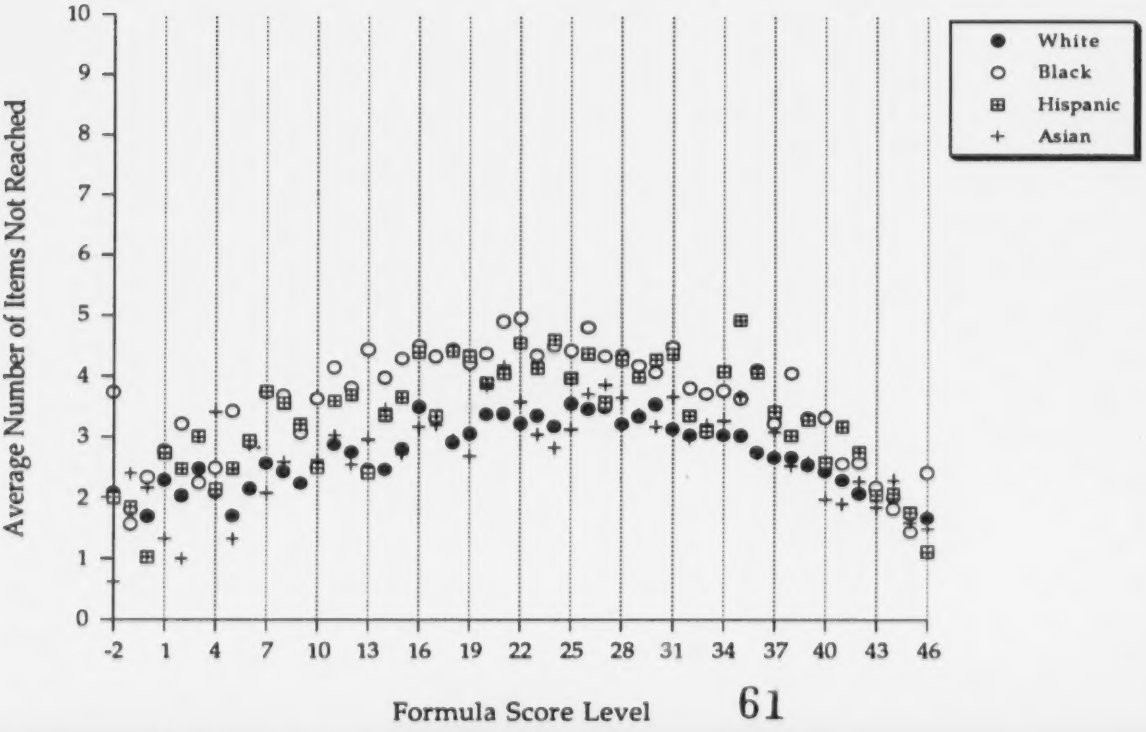


Figure 15: Form E Math by Gender

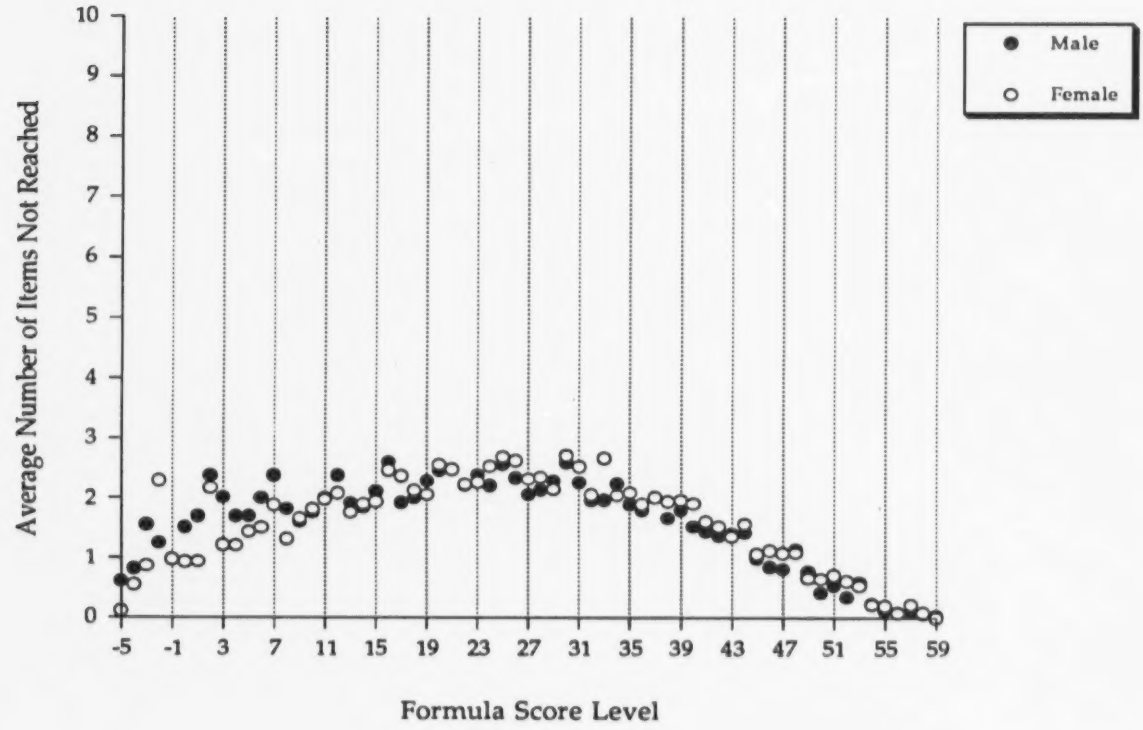


Figure 15: Form E Math by Race/Ethnicity

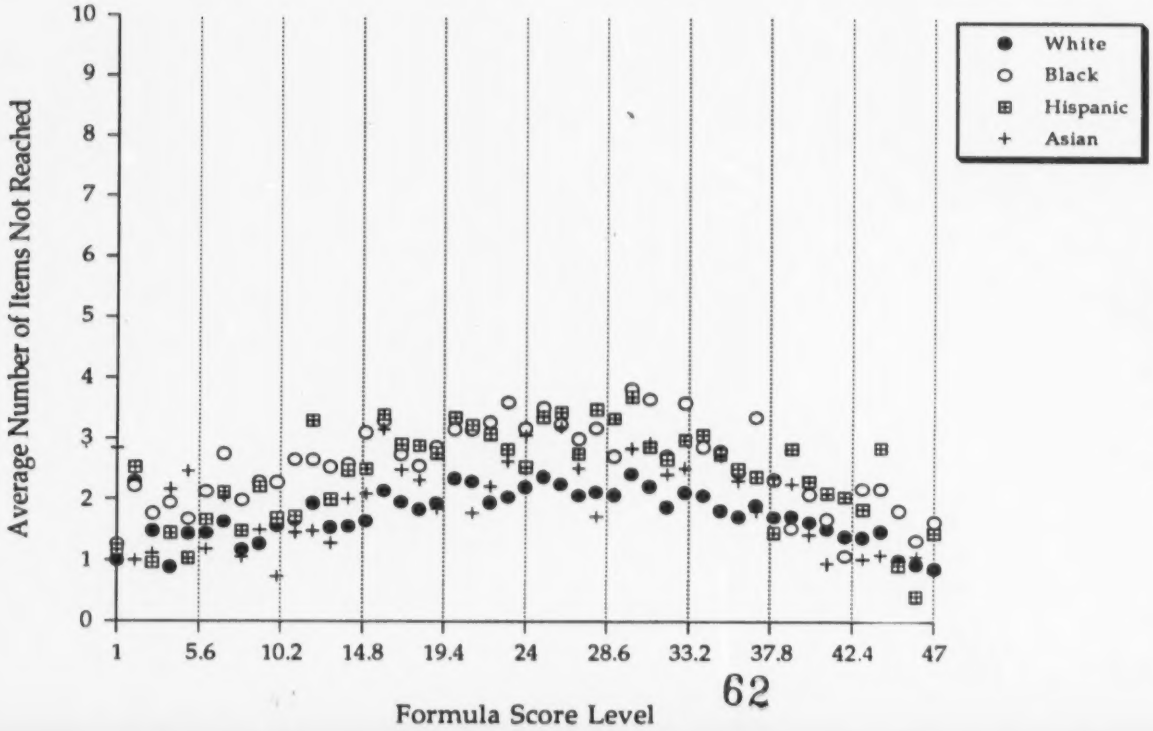


Figure 16: Form F Math by Gender

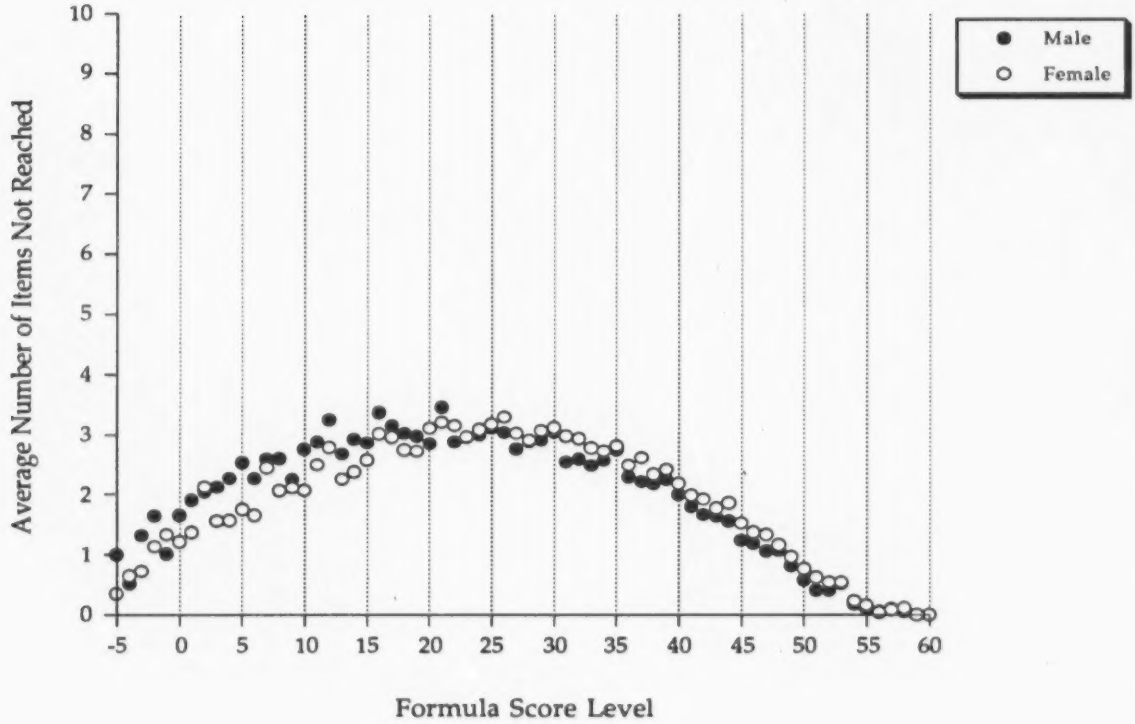


Figure 16: Form F Math by Race/Ethnicity

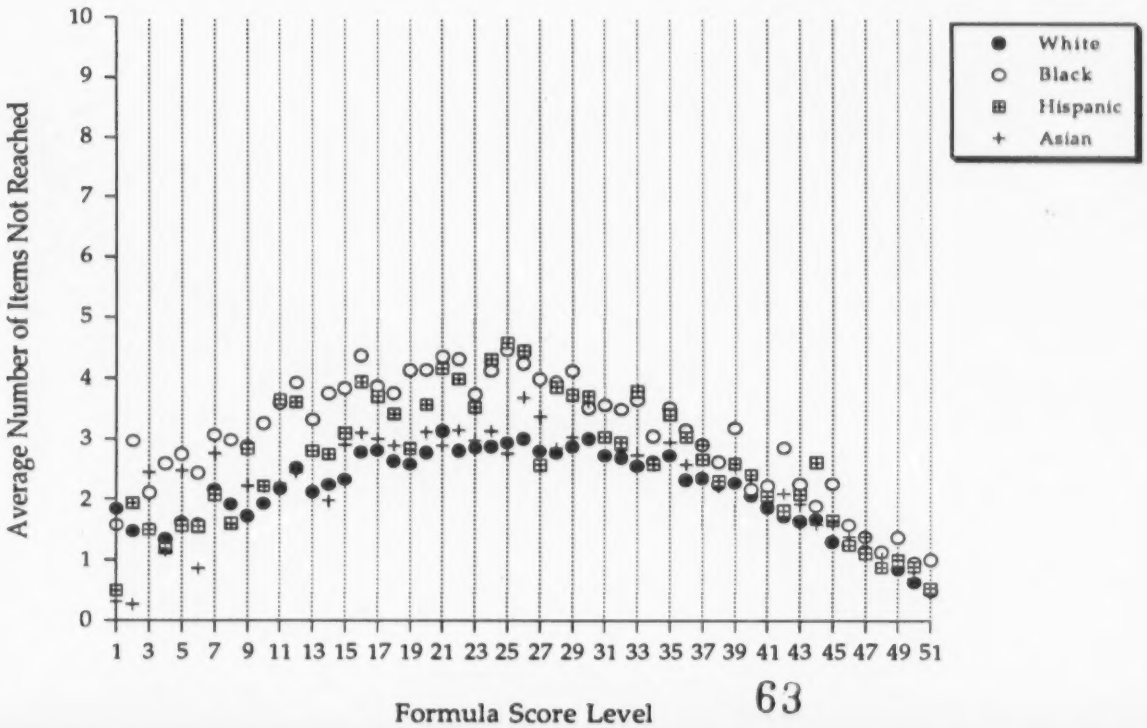


Figure 17: Form G Math by Gender

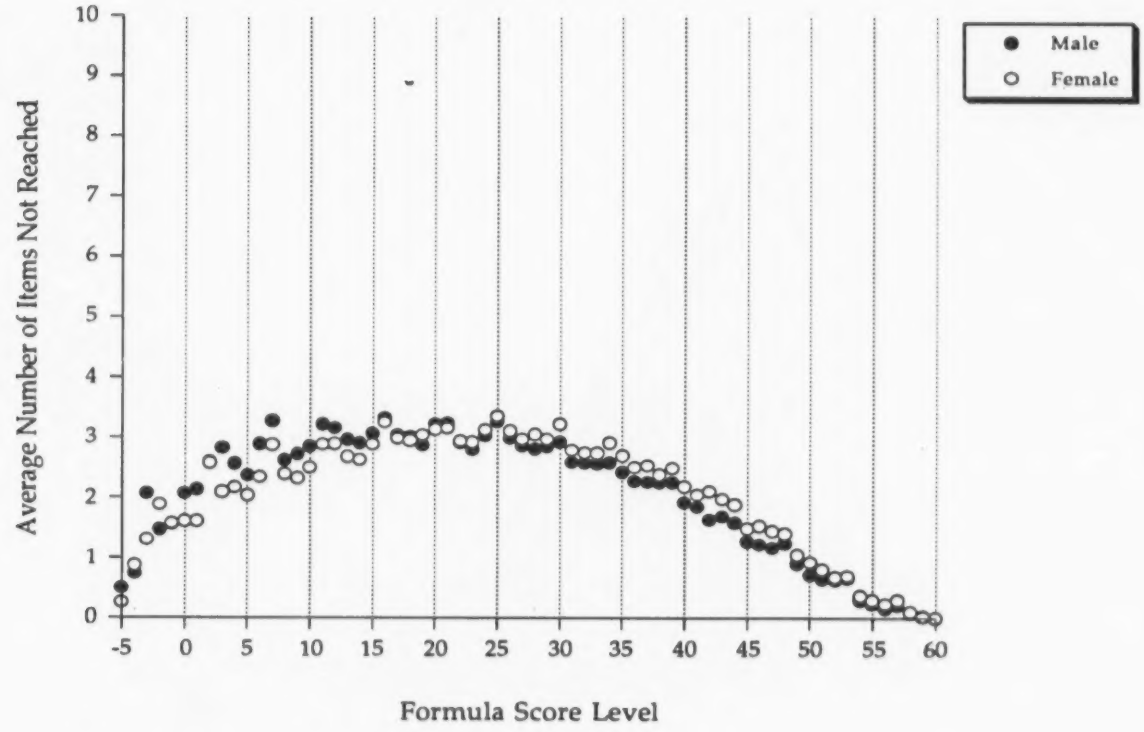


Figure 17: Form G Math by Race/Ethnicity

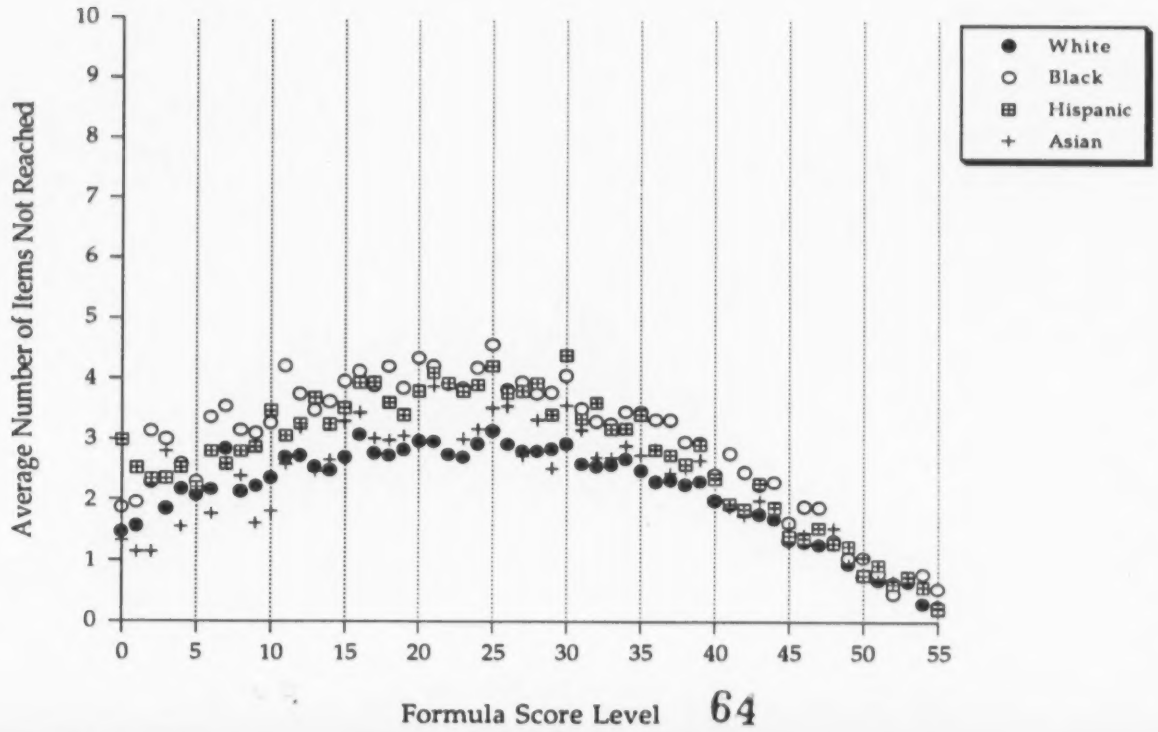


Figure 18: Form X Verbal by Gender

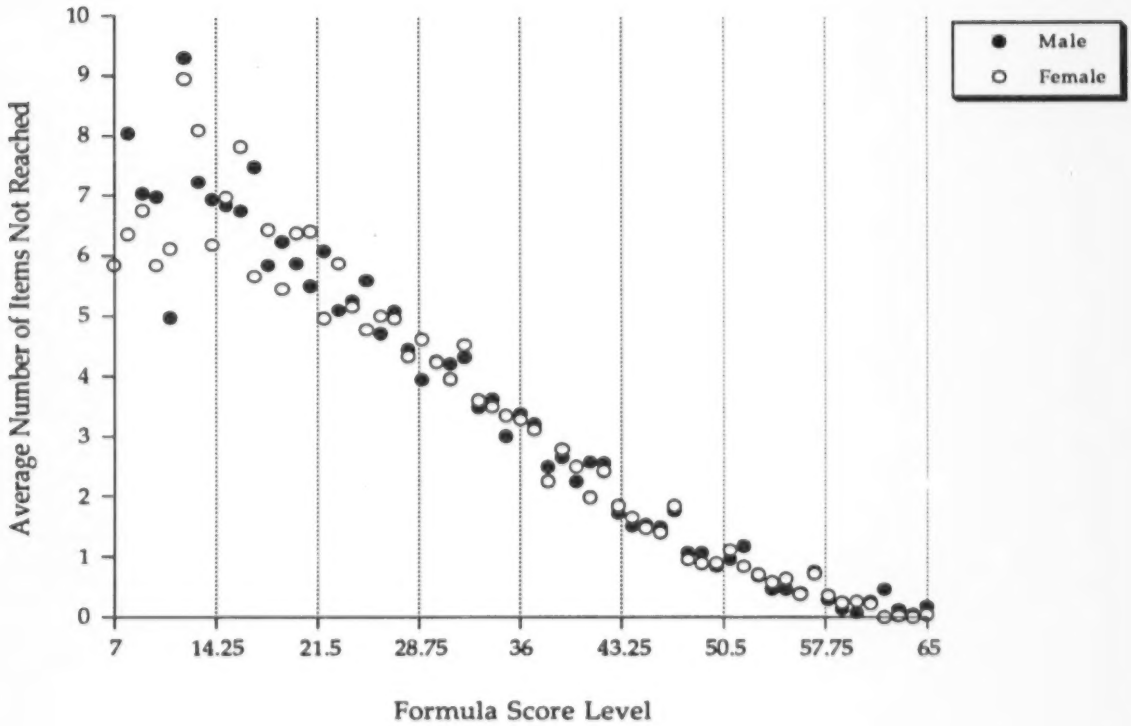


Figure 18: Form X Verbal by Race/Ethnicity

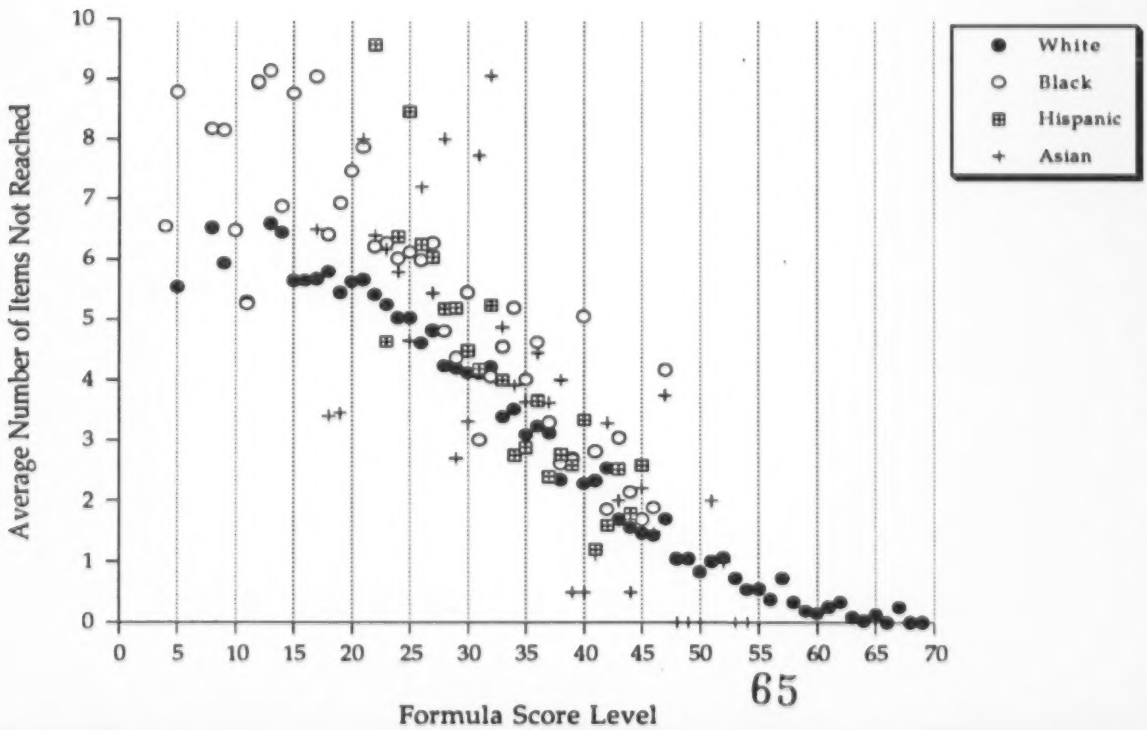


Figure 19: Form Y Verbal by Gender

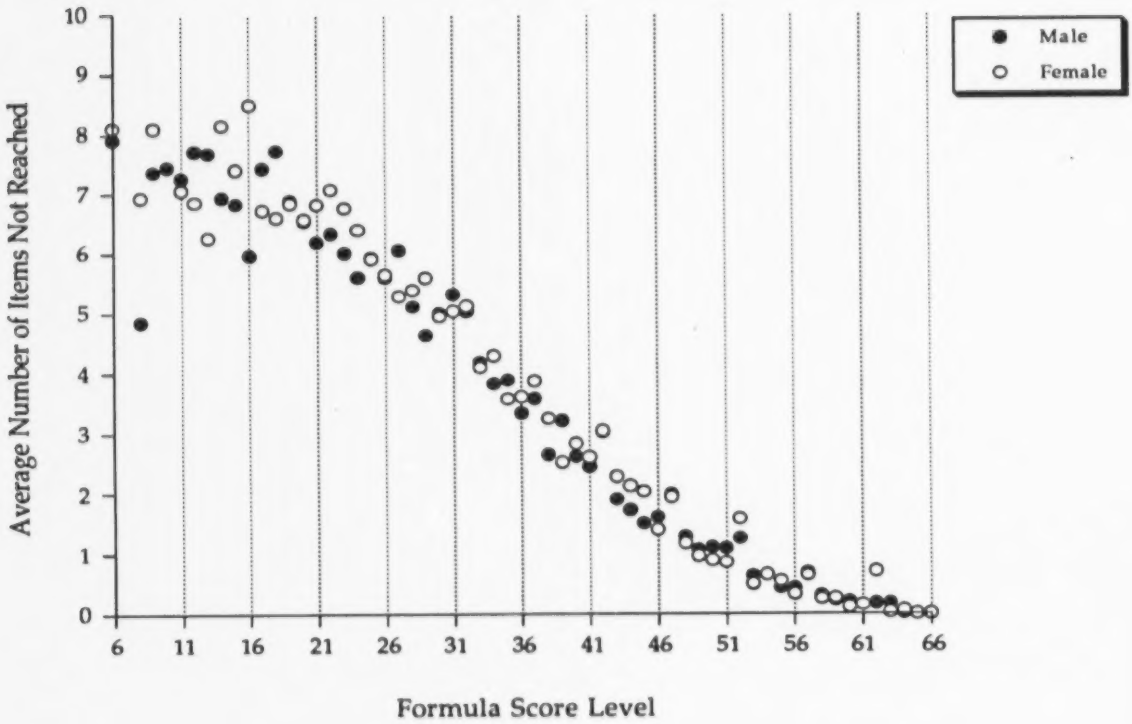


Figure 19: Form Y Verbal by Race/Ethnicity

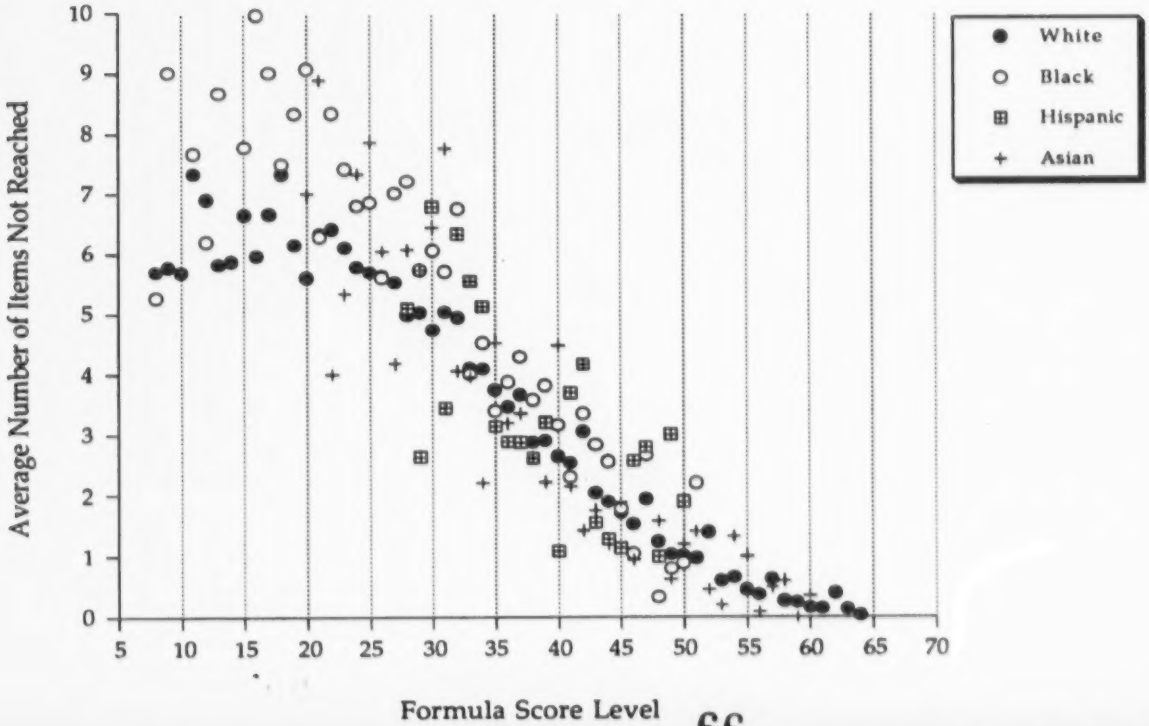


Figure 20: Form Z Verbal by Gender

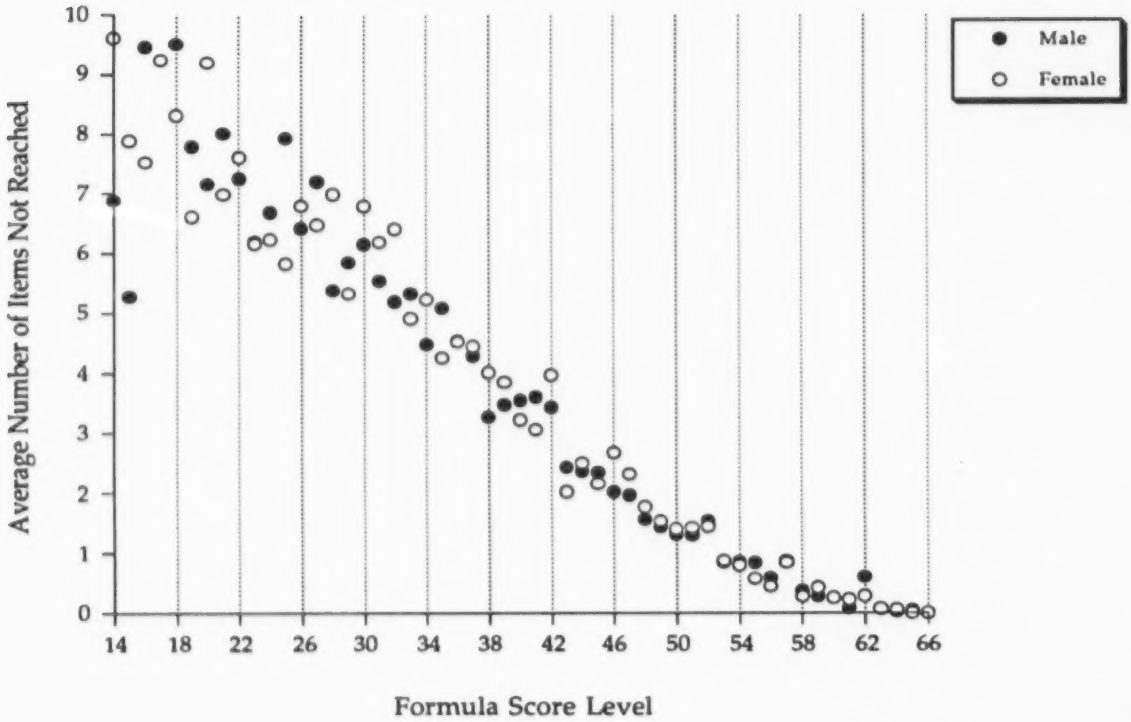


Figure 20 Form Z Verbal by Race/Ethnicity

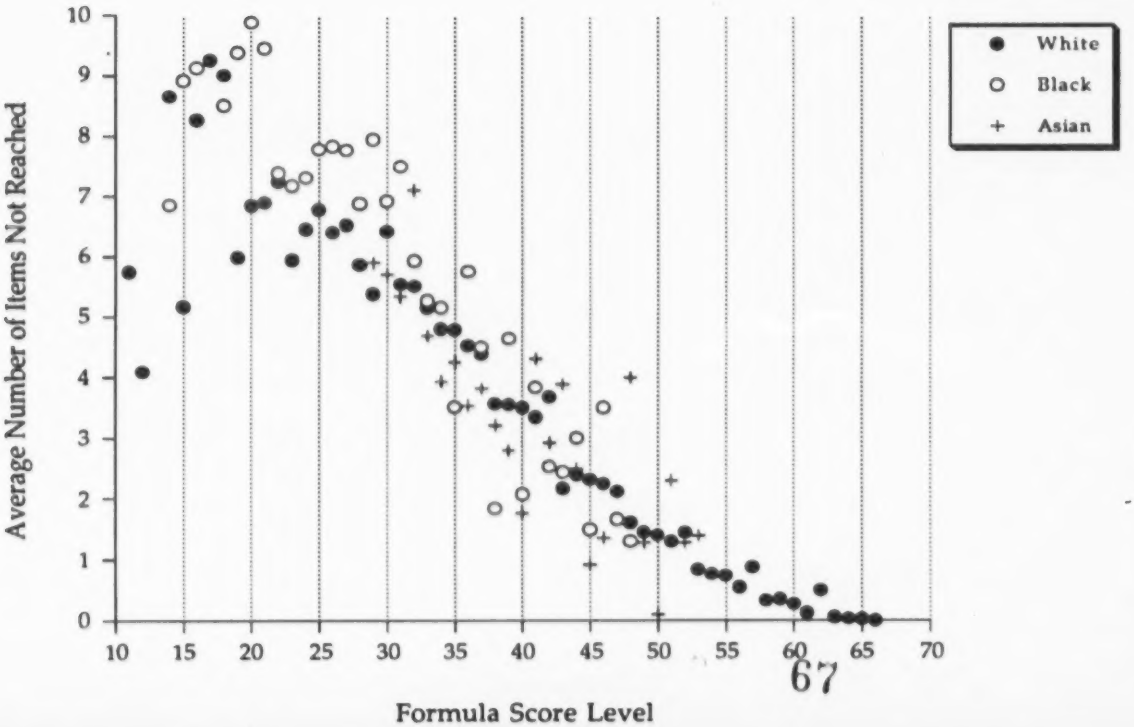


Figure 21: Form X Quantitative by Gender

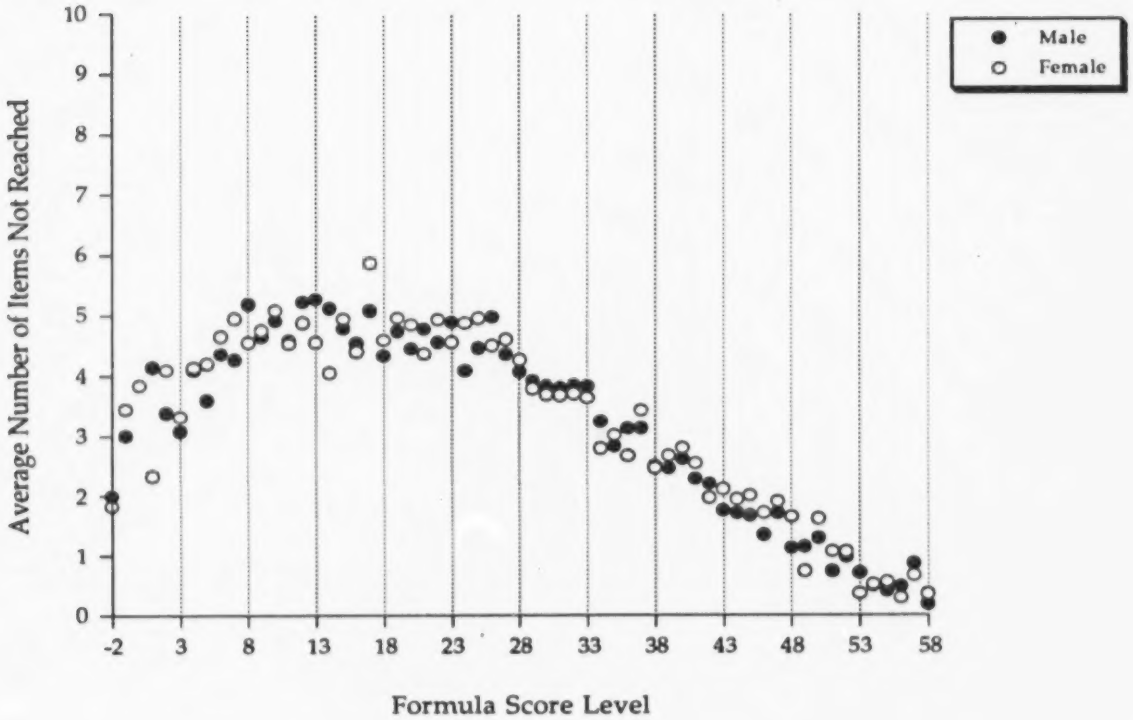


Figure 21: Form X Quantitative by Race/Ethnicity

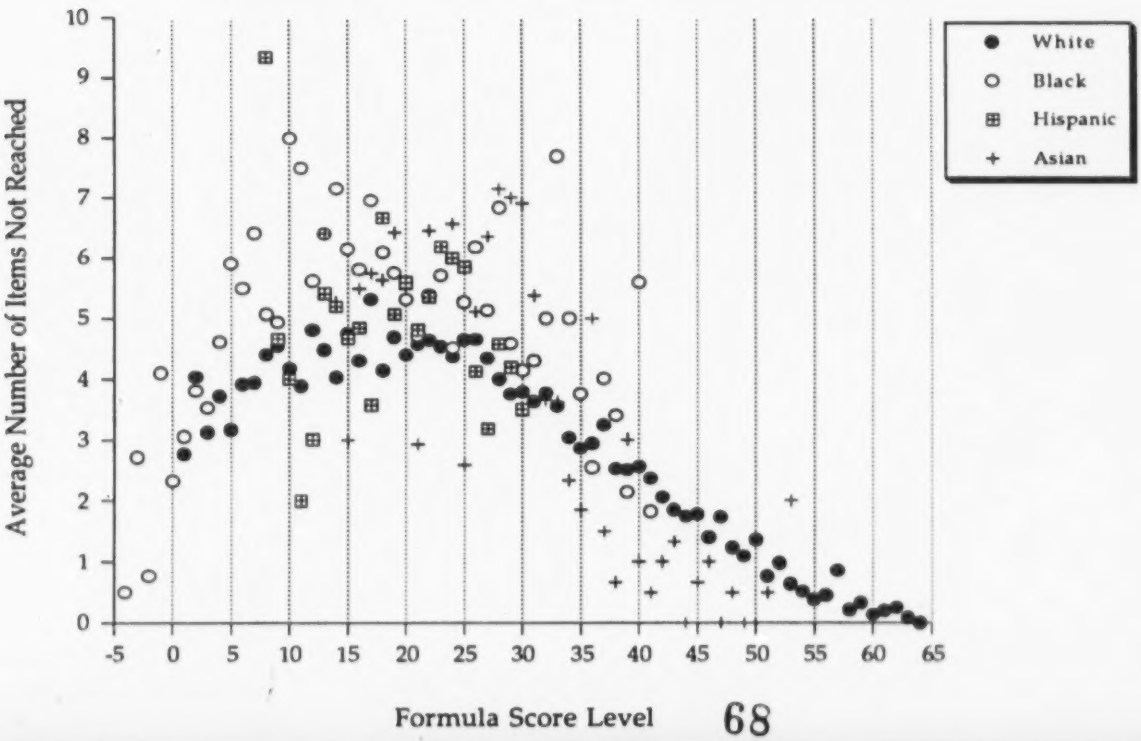


Figure 22: Form Y Quantitative by Gender

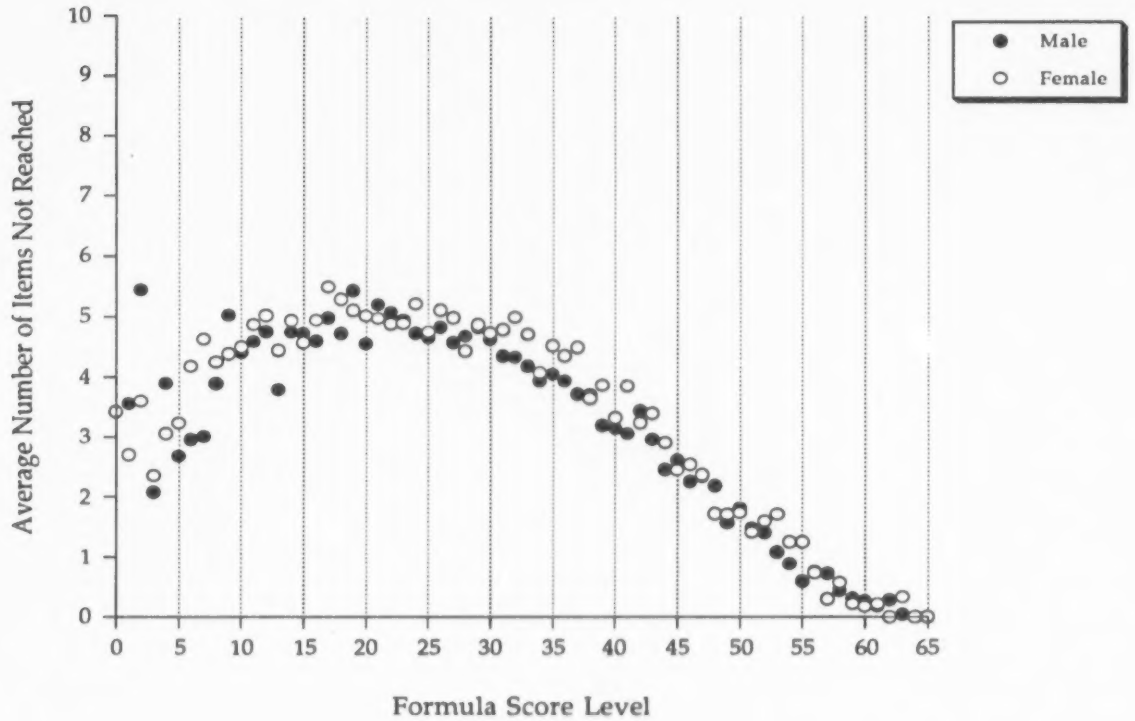


Figure 22: Form Y Quantitative by Race/Ethnicity

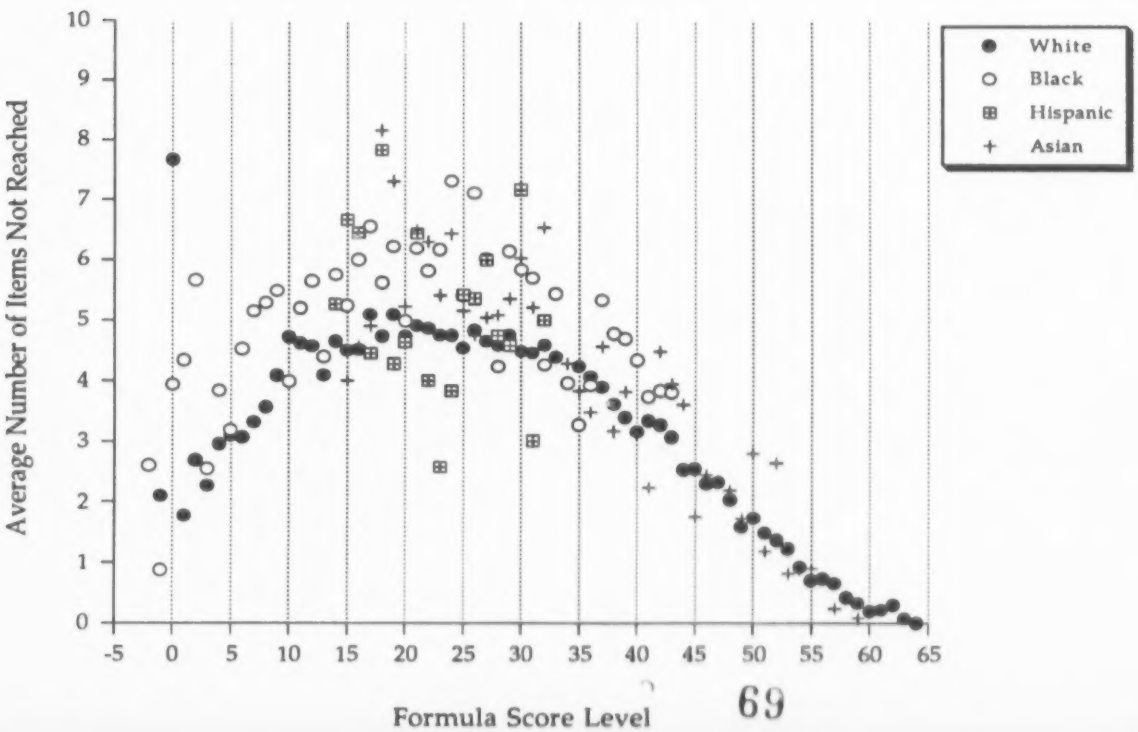


Figure 23: Form Z Quantitative by Gender

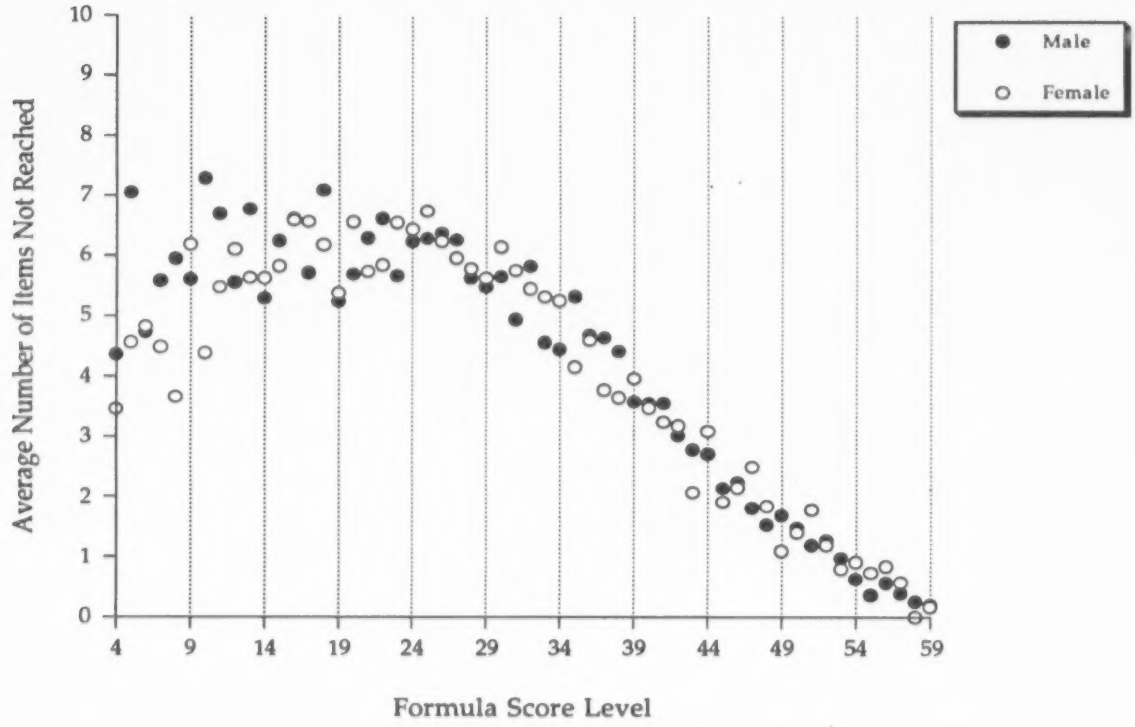


Figure 23: Form Z Quantitative by Race/Ethnicity

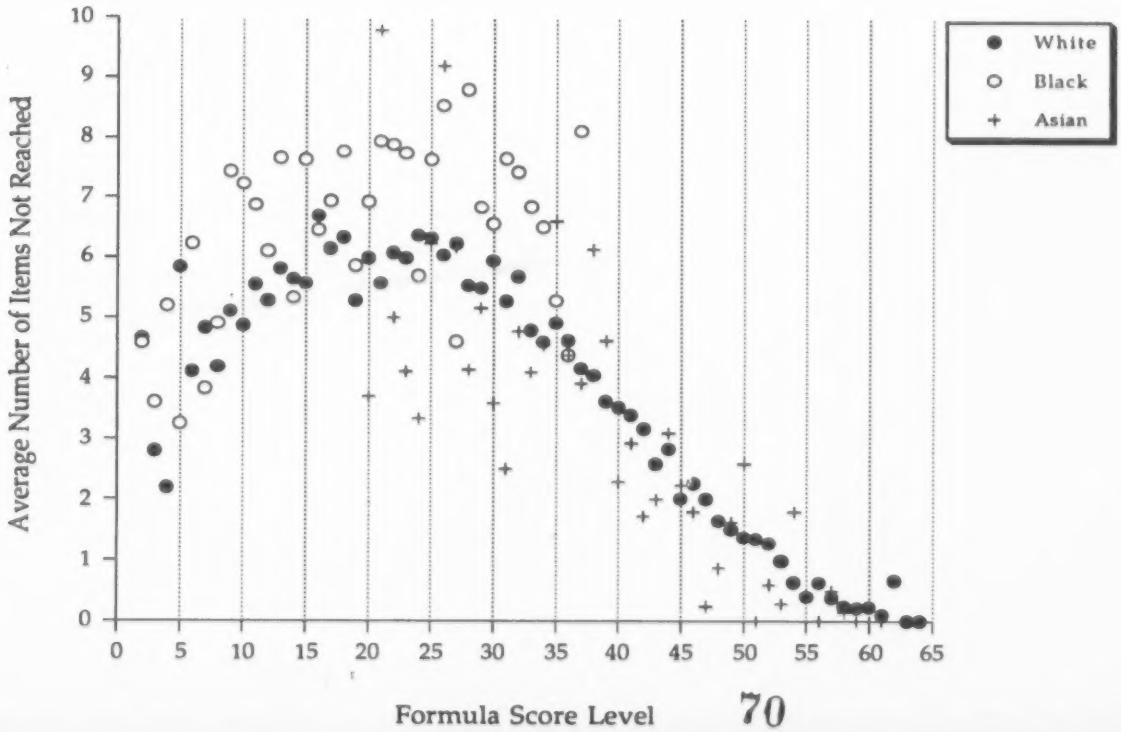


Figure 24

Standardized Differences in Percentages Omitting for Asian Americans and Whites, by Item Difficulty

(Form E Verbal)